Facultad de Óptica y Optometría
Universidad Complutense de Madrid

Doctoral Thesis

Development and Application of Bioimage Analysis Methods for Advanced Optical Microscopy

Ana Cayuela López

Development and Application of BioImage Analysis Methods for Advanced Optical Microscopy

Thesis by

Ana Cayuela López

Advisors

Carlos Óscar Sánchez Sorzano

José Antonio Gómez Pedrero

Facultad de Óptica y Optometría
Universidad Complutense de Madrid

This dissertation is submitted for the degree of
Doctor in Optics, Optometry and Vision

Madrid, October 2023

Desarrollo y Aplicación de Métodos de Análisis de Bioimágenes para Microscopía Óptica Avanzada

Presentada por

Ana Cayuela López

Directores

Carlos Óscar Sánchez Sorzano

José Antonio Gómez Pedrero

Facultad de Óptica y Optometría
Universidad Complutense de Madrid

Memoria presentada para optar al grado de
Doctor en Óptica, Optometría y Visión

Madrid, Octubre 2023

UNIVERSIDAD
COMPLUTENSE
MADRID

DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS PRESENTADA PARA OBTENER EL TÍTULO DE DOCTOR

D./Dña. Ana Cayuela López,
estudiante en el Programa de Doctorado Óptica, Optometría y Visión,
de la Facultad de Óptica y Optometría de la Universidad Complutense de Madrid, como autor/a de la tesis presentada para la obtención del título de Doctor y titulada:

Desarrollo y Aplicación de Métodos de Análisis de Bioimágenes para Microscopía Óptica Avanzada

Development and Application of BioImage Analysis Methods for Advanced Optical Microscopy

y dirigida por: Carlos Óscar Sánchez Sorzano y José Antonio Gómez Pedrero

DECLARO QUE:

La tesis es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, de acuerdo con el ordenamiento jurídico vigente, en particular, la Ley de Propiedad Intelectual (R.D. legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, modificado por la Ley 2/2019, de 1 de marzo, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), en particular, las disposiciones referidas al derecho de cita.

Del mismo modo, asumo frente a la Universidad cualquier responsabilidad que pudiera derivarse de la autoría o falta de originalidad del contenido de la tesis presentada de conformidad con el ordenamiento jurídico vigente.

En Madrid, a 16 de Octubre de 2023

Fdo.: ______________________________

Esta DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD debe ser insertada en la primera página de la tesis presentada para la obtención del título de Doctor.

Dissertation submitted by Ana Cayuela López to the *Faculty of Optics and Optometry of the Universidad Complutense de Madrid* in Partial Fulfillment of the Requirements for the Degree of *Doctor por la Universidad Complutense de Madrid en el Programa de Doctorado de Óptica, Optometría y Visión* .

Madrid, 2023.

Title:

Development and Application of Bioimage Analysis Methods for Advanced Optical Microscopy

PhD Student:

Ana Cayuela López (acayuela@cnio.es)
Biocomputing Unit
National Centre for Biotechnology
Spanish National Research Council
Cantoblanco, Madrid, Spain

Advisors:

Carlos Óscar Sorzano Sánchez (coss@cnb.csic.es)
José Antonio Gómez Pedrero (jagomezp@fis.ucm.es)

This research was supported by the Spanish MICINN `(PRE2018-086112)` by the FPI fellowship from the Spanish Ministry of Science and Innovation through the Severo Ochoa excellence accreditation `SEV-2017-0712- 18-1`.

To my dedicated and loving parents,
your unwavering commitment to equality and solidarity,
has shaped my values and inspired my journey…

Yo no sé lo que es el destino.
Caminando fui lo que fui…

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others. All sources, references, and literature used or excerpted during the elaboration of this work are properly cited and listed in complete reference to the due source.

Ana Cayuela López

Madrid, October 2023

Acknowledgements

Today is a special day, a day when I can finally say, "*It's over, because I set my mind to it and suffered, yes, I suffered a little!*" (Although it wasn't really that bad!) (Warning: it is a spanish song). After four years of immersion in the fascinating and sometimes complicated world of bioimage analysis at the National Centre for Biotechnology of CSIC, I have finally reached the final chapter of this thesis. This journey has been like a roller coaster of code lines and optical microscopy, and I have traveled it with a network of support and camaraderie that deserves a special moment of recognition to express my gratitude.

First and foremost, to my thesis supervisor, **Carlos Oscar**. "Rome wasn't built in one day," you used to say when my own confidence was at its lowest point. Your generosity in giving me this opportunity has changed my life in unimaginable ways, making me a better professional, and perhaps, a better person. To **José Antonio**, your wise advice and willingness to help have kept me on the right path.

Now, let's talk about my colleagues in the Biocomputing Unit at CNB, the people from **"B13"** (because, naming all of you would be impossible!). In this group, I have not only found a place to learn from the best, but it has also become a home, a true home where I can laugh endlessly and beyond! In B13, I have learned to accept myself as I am, sometimes beyond my capabilities. Here, the id, ego, and superego coexist harmoniously every day. There will always be someone in B13 more dissident than you, coming up with a bolder "idea" than yours, and of course, without fear of anything! Thank you for enriching my life in ways that go far beyond science. You have been the dream group on this journey, proving that friendship and science can be an explosive combination. With all of you, everything is possible!

To the wonderful people from the Correlative Microscopy Unit at CNB, **Jonathan, Javi, and David**. You are the best in Spanien! (And I mean it). Thank you for valuing me, sharing your knowledge, and making me feel like a part of your indomitable "dream team" of correlatives. You are inspiring, and I say that with all my heart! To the great colleagues in the Advanced Light Microscopy Unit at CNB, **Ana Oña**, **Gianluca**, and **Jaime**. Your constant availability and good vibes have rescued me on more than one occasion when I was lost. Thank you so much for making every day in the unit a true luxury!

And to the fourth floor of CNB and its people from the Department of Immunology and Oncology, **Jose, Anabel, Ester, Pilar, Dani, Vicente, Alberto, Miriam, Marina...** Even though I'm not officially part of it, well, I feel like I am! You've made me feel at home, or at least like I'm at a constant party in the hallway. I want to shout from the rooftops: "*Long live the fourth floor, damn it!*"

To my beloved and ever-needy group of the **"**RROH"**, who are, of course, also from the fourth floor, it couldn't be otherwise! **Sofi, Adri, Diego, and Cris**, some people say laughter is the best medicine, and you, you are medicine. In these pandemic years at CNB, I came across such a unique group that has pushed my sanity to its limits on more than one occasion. Damn you! I wouldn't change a second of what I've experienced, and I can't ask for friends more special than you. And to you, **Sofi**, what plugins have united, let no one separate. Thank you for bringing this gang of characters together. You've always been so good at gathering wonderful and dissident people! You're the reason we're here, and I can't help but be grateful for that. Let's keep laughing and facing the challenges that life and science have in store for us.

To my new colleagues in the Confocal Microscopy Unit at CNIO, **Isabel, Manu, and Yisus**. I want to thank you for being so welcoming in this new stage! Since I arrived at the unit, you have made me feel valued and at home, and that is priceless. I'm looking forward to seeing what the future holds for us and what we will achieve together.

To my tireless friends from Totana, my **Yai, Tonino, Terry, Dogo, Juancar, Marisa, Miris, Anicus, Esmeralda and Juan, Merry...** Even though our lives have taken different paths since those days in the "parque azul" endless afternoons of "manchados a un euro" at Tumar, and never-ending nights at the "Charca," we share a friendship that has withstood the test of time like a treasure buried in the streets of our denigrated yet beloved Totana. Despite the distance that separates us, for a long time now, I have never felt alone. You are always there, on the other end of the phone and in my memories, and those laughs (oh, the laughs we've had!) are etched into my skin every day. You are a part of me.

And to my colleagues from the **"El Plajo"** group in Totana, who would have thought in 2011 that "Cayu" would become a doctor? From our morning excursions to the "pinos", the fearsome escapes from the "Arco" and the afternoons on the "Tierno" wall hunting "micras", or rather avoiding being hunted, until our endless and "eventful" nights on the terrace of the "Aledo", each of you is part of one of the most beautiful chapters of my life

To **Carmen**, my eternal roommate. You have been a constant presence in my life, a friend and confidant in the darkest moments as well as the brightest. You came into my life when I was a young woman, gripped by my fear and insecurities, and you were there to lend

a helping hand and guide me through this journey of self-discovery and growth. Thank you for believing in me when I sometimes doubted myself.

To my entire family, and especially my parents, **Consuelo and Alfonso**. You are the true guiding light in my life! Your unconditional love, dedication, and mutual support are inspiring and drive every step I take on this journey called life. Your values are the foundation of my path, always reminding me of the importance of justice, equality, and social commitment. You give meaning to my life.

To you, my sister **María**, thank you for being my unwavering rock. You are more than a sister; you are my confidant, my companion, and above all, my friend. Your presence in my life is a treasure beyond words. Despite everything (and when I say 'everything,' I mean everything, you know...), I know I can always count on you. Let's continue facing life together.

To my dear **tata María**, my second mother, and her **"Pochili"**, who have sheltered and raised me since I was a baby, and whom I love with all my heart. Tata, you have cared for me with absolutely unconditional love. The tenderness and kindness you have shared with me guide me every day. And Pochili, your popular wisdom from a life lived and worked without choice has been essential in shaping my values. You are an essential part of who I am, and I carry you with me every day.

And to the rest of my beloved family, my **tita Mari** and **tita Paqui**, the adopted aunts from the 2004 and 2006 classes of Tierno Galván School in Totana, who were eagerly awaited by everyone. I hope to return half of the kindness, joy, and affection you have given me over the years. My **tita Paqui from the "Charca"**, an example of generosity and unconditional support. Always ready to give without expecting anything in return, your constant presence is a gift in my life. My **tito Ginés**, who was already a geek in a time when being one required strict discipline, dedication, and subscriptions to "Discoplay" journal. I deeply appreciate your countercurrent tastes and your constant positive influence on my love for music and movies.

To those who are no longer with us, my grandfather **Ginés "el del pan"** and my grandmother **Kiska "la del horno"**, my grandmother **Maria Jesús**, and my grandfather **Tomás "el frascuelo"**. To you, who couldn't choose and had to face life's toughest circumstances, know that your granddaughter is now a doctor. To **tía Rosario** and **tía "Rojica"**, because they were, we are. Because we are, they will be. Your memory and love fill my heart.

And to you, my **Guillermo**, my "feti," you are proof that love and humor can go hand in hand in the most beautiful way. On this journey, you have been my accomplice and refuge. Thank you for filling my world with joy, tenderness, and adding that touch of absurdity to every day. Let's continue building moments that we will treasure forever.

Agradecimientos

Hoy es un día especial, un día en el que finalmente puedo decir *"¡se acabó, porque yo me lo propuse y sufrí, y sí, sufrí un poquito!"* (¡Aunque tampoco fue para tanto!). Después de cuatro años de inmersión en el fascinante y a veces complicado mundo del análisis de bioimágenes en el Centro Nacional de Biotecnología del CSIC, finalmente he llegado al capitulo final de esta tesis. Este viaje ha sido como una montaña rusa de líneas de código y microscopía óptica, y lo he recorrido con una red de apoyo y camaradería que merece un momento especial de reconocimiento para expresar mi gratitud.

En primer lugar, a mi supervisor de tesis, **Carlos Oscar**. *"Roma no se construyó en un día"*, me repetías, cuando mi propia confianza en mí misma estaba en su punto más bajo. Tu generosidad al darme esta oportunidad ha cambiado mi vida de manera inimaginable, convertiendome en una mejor profesional y, quizás, en una mejor persona. A **José Antonio**, tus sabios consejos y disposición me han mantenido en el rumbo correcto.

Ahora, hablemos de mis compis de la Unidad de Biocomputación del CNB, vamos las gentes del **"B13"** (y es que, ¡sería imposible nombraros a tod@s!). En este grupo no solo he encontrado un lugar para aprender de los y las mejores, sino que se ha convertido en hogar, un verdadero hogar donde poder partirme de la risa hasta el infinito, ¡y más allá!. En en B13, he aprendido a aceptarme tal y como soy, a veces por encima de mis posibilidades. Aquí el ello, el yo y el superyó coexisten en armonía a diario. Siempre habrá alguien en el B13 más disidente que tú, soltando una "ocurrencia" más audaz que la tuya, ¡y por supuesto, sin miedo a nada!. Gracias por enriquecer mi vida de maneras que van mucho más allá de la ciencia. Habeis sido el grupo soñado en esta travesía, demostrando que la amistad y la ciencia pueden ser una combinación explosiva.¡Con vosotros y vosotras todo es posible!

A las buenas gentes de la Unidad de Microscopía Correlativa del CNB, **Jonathan, Javi y David**. ¡Sois la caña de Spanien! (y de verdad de la buena). Gracias por valorarme, por compartir vuestro conocimiento y por hacerme sentir parte de vuestro irreductible "dream team" de correlativos. Sois inspiradores ¡y lo digo con todo mi corazón!

A los geniales compis de la Unidad de Microscopía Óptica Avanzada del CNB, **Ana Oña**, **Gianluca** y **Jaime**. Vuestra siempre disposición y buen rollo, me han rescatado en más de una ocasión cuando perdía el norte. ¡Mil gracias por hacer que cada día en la unidad sea un auténtico lujo!

Y a la cuarta planta del CNB y sus gentes del Departamento de Inmunología y Oncología, **Jose, Anabel, Ester, Pilar, Dani, Vicente, Alberto, Miriam, Marina...** a pesar de no ser oficialmente parte de ella, ¡pues sí, también me siento de allí!. Me habéis hecho sentir como en casa, o al menos como si estuviera en una fiesta constante en el pasillo. Clamorosamente, quiero gritar a los cuatro vientos: *¡Viva la cuarta planta, carajo!*

A mi amado y siempre necesitado, grupo de los **"**RROH"** que por supuesto, también son de la cuarta planta, ¡no podía ser de otra manera!. **Sofi, Adri, Diego y Cris**, dicen que la risa es el mejor remedio, y vosotros, es que sois medicina. En estos años pandémicos en el CNB, me topé con un grupo tan, pero tan singular, que en más de una ocasión ha llevado mi cordura al límite. ¡Malditos seáis!. No cambiaría ni un segundo de lo vivido, y no puedo pedir amigos más especiales que vosotros. Y a ti, **Sofi**, que lo que han unido los plugins, no lo separe el hombre. Gracias por traer a esta pandilla de personajos juntos.¡Siempre se te ha dado tan bien reunir a personas maravillosas a la par que disidentes!. Eres la culpable de que estemos aquí, y no puedo evitar estar agradecida por ello. Sigamos riéndo, y enfrentando los desafíos que la vida y la ciencia nos depara.

A mis nuevos compis de la Unidad de Microscopía Confocal del CNIO, **Isabel, Manu y Yisus**. ¡Quiero daros las gracias por ser tan acogedores en esta nueva etapa! Desde que llegué a la unidad, me habéis hecho sentir muy valorada y como en casa, y eso no tiene precio. Estoy deseando ver lo que el futuro nos tiene preparado y todo lo que lograremos juntos.

A mis incombustibles totaneros, mi **Yai, Tonino, Terry, Dogo, Juancar, Marisa, Miris, Anicus, Esmeralda y Juan, Merry...** A pesar de que nuestras vidas han tomado caminos diferentes, desde aquellos días en el "parque Azul", tardes infinitas de "manchados de euro" en el Tumar y noches infinitas en la "Charca", compartimos una amistad que ha resistido el paso del tiempo como un tesoro enterrado en las calles de nuestra denostada, y siempre amada, Totana. Aunque la distancia nos separa, ya desde hace mucho tiempo, nunca me he sentido sola, siempre estais ahí, al otro lado del teleéfono, y en mis recuerdos, y esas risas (¡ay qué risas nos hemos echado!) están forjadas en mi piel a diario. Sois parte de mi.

Y a mis colegas de la **peña "El Plajo"** de Totana, ¿quién nos iba a decir en 2011 que la "Cayu" sería doctora? Desde nuestras excursiones mañaneras a los "pinos", las temibles huídas del "Arco" y las tardes en el muro del "Tierno" cazando "micras", o más bien evitando ser cazados, hasta nuestras noches interminables y "accidentadas" en la terraza del "Aledo", cada uno de vosotros forma parte de uno de los capítulos más hermosos de mi vida.

A **Carmen**, mi eterna compi de piso. Has sido una presencia constante en mi vida, una amiga y confidente en los momentos más oscuros y también en los más luminosos.

Llegaste a mi vida cuando era una joven presa de mi pánico e inseguridades, y tú estuviste allí para tenderme una mano amiga y guiarme a lo largo de este viaje de autodescubrimiento y crecimiento. Gracias por creer en mí cuando yo misma a veces dudaba.

A toda mi familia, y en especial a mis papis, **Consuelo y Alfonso**. ¡Vosotros sí que sois el verdadero faro en mi vida! Vuestro amor incondicional, dedicación y ejemplo de apoyo mutuo, inspiran e impulsan cada paso dado en este viaje que es la vida. Vuestros valores son el cimiento de mi camino, recordándome siempre la importancia de la justicia, la igualdad y el compromiso social. Dais sentido a mi vida.

A ti, mi hermana **María**, gracias por ser mi roca inquebrantable. Eres mucho más que una hermana, eres mi confidente, mi compañera y, sobre todo, mi amiga. Tu presencia en mi vida es un tesoro más allá de las palabras. A pesar de todo, (y cuando digo 'todo', es todo, ya tu sabes...), sé que siempre puedo contar contigo. Sigamos enfrentando la vida juntas.

A mi querida **tata María**, mi segunda madre, y su **"Pochili"**, quienes me han arropado y criado desde bebé, y a los que quiero con todo mi corazón. Tata, has cuidado de mí con un amor absolutamente incondicional. La ternura y bondad que has compartido conmigo me guían cada día. Y Pochili, tu sabiduría popular del que ha vivido y trabajado sin elección, han sido fundamentales para forjar mis valores. Sois una parte esencial de quien soy, y cada día os llevo conmigo.

Y al resto de mi querida familia, mi **tita Mari** y **tita Paqui**, las por todo el mundo ansiadas titas adoptivas de las promociones del 2004 y 2006 del colegio Tierno Galván de Totana. Ojalá devolveros la mitad de la bondad, alegría y cariño que me habeis dado a lo largo de los años. Mi **tita Paqui de la "Charca"**, un ejemplo de generosidad y apoyo incondicional. Siempre dispuesta a dar sin esperar nada a cambio, tu presencia constante es un regalo en mi vida. Mi **tito Ginés** que ya era friki en una época donde serlo requería de alta disciplina, dedicación y suscripciones a la "Discoplay". Aprecio profundamente tus gustos a contracorriente y tu constante influencia positiva en mi amor por la música y el cine.

A los que ya no están, mi yayo **Ginés "el del pan"** y mi **Kiska "la del horno"**, a mi yaya **Maria Jesús** y mi yayo **Tomás "el frascuelo"**. A vosotros que no pudisteis elegir y que tuvisteis que hacer frente a las circunstancias más duras de la vida, que sepáis que vuestra nieta es doctora. A la **tía Rosario** y la **tía "Rojica"**, porque fueron, somos. Porque somos, serán. Vuestro recuerdo y amor llenan mi corazón.

Y a ti, mi **Guillermo**, mi "feti", eres la prueba de que el amor y el humor pueden ir de la mano de la manera más hermosa. En este viaje, has sido cómplice y refugio. Gracias por llenar mi mundo de alegría, ternura y ponerle a todos los días ese toque de lo absurdo. Sigamos construyendo momentos que atesoraremos para siempre.

Abstract

The field of biological image analysis has undergone significant transformations over the past few decades due to advancements in microscopy techniques and computer vision. The journey of biological image processing and subesequent analysis started with the transition from traditional photographic imaging to digital acquisition, enabling the development of computer-based algorithms for noise reduction, contrast enhancement, and object counting. From the 1970*s* to the 1990*s*, fundamental algorithms for image segmentation, object recognition, and feature extraction were established, laying the foundation for quantitative analysis of cellular structures. The introduction of machine learning and artificial intelligence approaches into biological image analysis is a significant turning point, leading to high-throughput analysis of large-scale microscopy datasets. The synergistic integration of computer vision and biology fields has given rise to bioimage analysis, a multidisciplinary field for extracting quantitative information of images acquired from biological samples for further obtain meaningful biological insights. Bioimage analysis deals with increasing volumes of microscopy data generated by current high-throughput microscopy modalities. Therefore, the emergence of interdisciplinary collaborations between biologists and computer scientists led to the development of specialized tools and platforms for image analysis and data management.

The primary objectives of this thesis revolve around extending the customization and automation in the quantitative analysis of fluorescence microscopy images (single-particle tracking, image registration, cell-type classification...) at the National Centre for Biotechnology (CNB). Hence with the establishment of a Quantitative Image Analysis Unit (QIAU), we aimed to leverage cutting-edge microscopy facilities and advance quantitative biology and bioimaging power at CNB. The thesis also emphasizes the transition from qualitative to quantitative analysis, involving the development of tools to extract quantitative information from large image datasets, and implementing real-time image processing. The thesis presents several open-source tools to address these objectives. `Cell-TypeAnalyzer`, a Fiji plugin which facilitates the user-defined classification of specific cell types based on morphological, intensity, and spatial features. `TrackAnalyzer` which focuses on single-particle tracking (SPT) analysis, providing a user-friendly interface for customizable SPT analyses, including spot detection, trajectory reconstruction, or diffusion and motion analysis. The thesis also

introduces the `OFM-Corrector` protocol, which offers real-time image registration to compensate for geometric distortions in fluorescence microscopy images. These tools aimed to enhance the accuracy, reproducibility, and efficiency of bioimage analysis.

Overall, the thesis contributes to the evolution of automated and quantitative analysis in optical microscopy, with implications for understanding complex biological processes at the cellular level. Furthermore, the tools and methods presented in this thesis offer potential for further development and integration with existing microscopy platforms, paving the way for more efficient, accurate, and user-friendly bioimage analysis.

Keywords: bioimage analysis, optical microscopy, fluorescence microscopy, image processing, automation, quantitative analysis, cell-type, image registration, single-particle tracking, open-source, chromatic aberration, real-time.

Resumen

El campo del análisis de imágenes biológicas ha experimentado transformaciones significativas en las últimas décadas debido a los avances en las técnicas de microscopía y la visión por computadora. El procesamiento de imágenes biológicas y su posterior análisis comenzó con la transición de la imagen fotográfica tradicional a la adquisición digital, lo que permitió el desarrollo de algoritmos basados en computadora para la reducción de ruido, realce de contraste y conteo de objetos. Desde la década de 1970 hasta la de 1990, se establecieron algoritmos fundamentales para la segmentación de imágenes, el reconocimiento de objetos y la extracción de características, sentando las bases para el análisis cuantitativo de estructuras celulares. La introducción del aprendizaje automático e inteligencia artificial en el análisis de imágenes biológicas es un punto de inflexión sin precedente, lo que lleva al análisis de alto rendimiento de grandes conjuntos de datos de microscopía a gran escala. La simultánea integración de la visión por computadora y biología ha dado lugar al análisis de bioimágenes, un campo multidisciplinario para extraer información cuantitativa de imágenes adquiridas de muestras biológicas para obtener conocimientos biológicos significativos. El análisis de bioimágenes se ocupa del aumento de volúmenes de datos de microscopía generados por las actuales modalidades de microscopía de alto rendimiento. Por lo tanto, la aparición de colaboraciones interdisciplinarias entre biólogos y científicos de la computación condujo al desarrollo de herramientas y plataformas especializadas para el análisis de imágenes y la gestión de datos.

Los objetivos principales de esta tesis giran en torno a la ampliación de la personalización y la automatización en el análisis cuantitativo de imágenes de microscopía de fluorescencia (seguimiento de partículas individuales, registro de imágenes, clasificación de tipos de células...) en el Centro Nacional de Biotecnología (CNB). Por lo tanto, con la creación de una Unidad de Análisis Cuantitativo de Imágenes (QIAU), teníamos como objetivo aprovechar las instalaciones de microscopía de vanguardia y avanzar en la potencia de la biología cuantitativa y la bioimagen en el CNB. La tesis también enfatiza la transición de un análisis cualitativo a uno cuantitativo, que implica el desarrollo de herramientas para extraer información cuantitativa de grandes conjuntos de datos de imágenes e implementar el procesamiento de imágenes en tiempo real. La tesis presenta varias herramientas de código abierto para

abordar estos objetivos. `Cell-TypeAnalyzer`, un plugin de Fiji que facilita la clasificación definida por el usuario de tipos de células específicos basada en características morfológicas, de intensidad y espaciales. `TrackAnalyzer` se enfoca en el análisis de seguimiento de partículas individuales (SPT), proporcionando una interfaz fácil de usar para análisis de SPT personalizables, que incluyen la detección de puntos, la reconstrucción de trayectorias o el análisis de difusión y movimiento. La tesis también presenta el protocolo `OFM-Corrector`, que ofrece el registro de imágenes en tiempo real para compensar las distorsiones geométricas en las imágenes de microscopía de fluorescencia. Estas herramientas tienen como objetivo mejorar la precisión, la reproducibilidad y la eficiencia del análisis de bioimágenes.

En resumen, la tesis contribuye a la evolución del análisis automatizado y cuantitativo en la microscopía óptica, con implicaciones para la comprensión de procesos biológicos complejos a nivel celular. Además, las herramientas y métodos presentados en esta tesis ofrecen potencial para un desarrollo adicional e integración con las plataformas de microscopía existentes, allanando el camino para un análisis de bioimágenes más eficiente, preciso y fácil de usar.

Palabras clave: análisis de bioimágenes, microscopía óptica, microscopía de fluorescencia, procesamiento de imágenes, procesamiento de imágenes, automatización, análisis cuantitativo, tipo de célula, registro de imágenes, seguimiento de partículas individuales, código abierto, aberración cromática, tiempo real.

Table of Contents

List of Figures

List of Tables

Nomenclature

Acronyms / Abbreviations

ADC Analog to Digital Converter

ADC Analog-to-Digital Counts

ALMF Advanced Light Microscopy Facility

ANN Artificial Neural Network

AWS Amazon Web Services

CA Chromatic Aberration

CARE Content-Aware Image Restoration

CCD Charge-coupled Device

CCF Cross Correlation Function

CLSM Confocal Laser Scanning Microscopy

CM Correlative Microscopy

CMI Correlated Multi-modal Imaging

CNB National Centre for Biotechnlogy

CNNs Convolutional Neural Networks

SDCM Spining-Disk Confocal Microscopy

CSIC Spanish National Research Council

DCT Discrete Cosine Transform

DFT	Direct Fourier Transform
DL	Deep Learning
DT	Decision Trees
EMBL	European Molecular Biology Organization
EM	Electron Microscopy
ERC	European Research Council
FCCS	Fluorescence Cross-Correlation Spectroscopy
FFT	Fast Fourier Transform
FIB	Focused Ion Beam
FNN	Fully Conected Neural Network
FOV	Field of View
FT	Fourier Transform
GCP	Google Cloud Platform
GFP	Green Fluorescent Protein
GPU	Graphics Processing Unit
GQD	Graphene Quantum Dots
HPC	High Perfomance Computing
HR	High-Resolution
I2PC	Instruct Image Processing Centre
IFT	Inverse Fourier Transform
DT	International Symposium on Biomedical Imaging
KNN	k-Nearest Neighbors
LM	Light Microscopy
LoG	Laplacian of Gaussian

LSFM Light-Sheet Fluorescence Microscopy

MCC Manders Colocalization Coefficient

MHT Multiple-Hypothesis Tracking

ML Machine Learning

MOC Manders Overlap Coefficient

MPM Multiphoton Microscopy

MSD Mean Squared Displacement

MSS Moment Scaling Spectrum

MTF Modulation Transfer Function

NA Numerical Aperture

NAS Network-Attached Storage

NIR Near-Infrared

NN Neural Network

OSI Open Source Initiative

OSS Open Source Software

PALM Photo-Activated Localization Microscopy

PCC Pearson’s Correlation Coefficient

PMT Photomultiplier Tube

PSF Point Spread Function

QIAU Quantitative Image Analysis Unit

RANSAC Random Sample Consensus

RF Random Forests

SCT Single Cell Tracking

SEM Scanning Electron Microscopy

SE Stimulated Emission

SIFT Scale-Invariant Feature Transform

SML Single Molecule Localization

SNR Signal-to-Noise Ratio

SPT Single Particle Tracking

SR-SIM Super-Resolution Structured Illumination Microscopy

SRM Super-Resolution Microscopy

STED Stimulated Emission Depletion

STORM Stochastic Optical Reconstruction Microscopy

SURF Speeded Up Robust Features

SVM Support Vector Machines

SWMS Scientific Workflow Management Systems

TIRF Total Internal Reflection Fluorescence

TVR Total Variation Regularization

YOLO You-Only-Look-Once

Chapter 1

Introduction

Microscopy revolutionized our understanding of biological systems, exploring cells and tissues at cellular and molecular levels. Over the past 50 years, parallel advancements in imaging techniques and computational tools enabled researchers to extract detailed information and gain profound insights into complex biological processes [1]. The journey of image processing and biological image analysis began with the emergence of digital imaging technologies in the late 20^{th} century. This transition from traditional photographic imaging to digital acquisition, allowed researchers to capture, store and manipulate images digitally, leading to the exploration of computer-based algorithms for such as noise reduction, contrast enhancement, and image restoration. With digital imaging becoming more prevalent, rudimentary algorithms for intensity quantification and object counting were developed, specifically tailored for biological imaging and enabling quantitative analysis of microscopy images.

Early methods were labor-intensive and limited in scope, but from the 1970*s* to the 1990*s*, fundamental algorithms for image segmentation, object recognition and feature extraction were devised. These algorithms, based on mathematical models and statistical approaches, enabled the identification and quantification of cellular structures. Image registration algorithms were also developed to align images from different time points or modalities. Advances in computer vision, machine learning and statistical modeling further expanded the capabilities of image processing. Powerful algorithms now handle diverse imaging modalities, from brightfield and fluorescence microscopy to electron microscopy [2]. Additionally, tools expanded to encompass sophisticated quantitative algorithms to extract features of morphology, cellular dynamics, and molecular interactions. The simultaneous integration of multiple microscopy modalities, provided a comprehensive understanding of biological systems, enabling researchers to probe cellular structures and functions from different perspectives.

The turn of the new millennium, witnessed the integration of machine learning and computational intelligence approaches into the field of biological image analysis. Techniques such as neural networks or random forests were included to automate image analysis, reducing manual intervention and subjective interpretations [3]. These methods facilitated the recognition of complex patterns, enabling high-throughput analysis of large-scale microscopy datasets [4]. In the early 2000*s*, with the explosive growth of biological data, the integration of bioinformatics and bioimage analysis became crucial for extracting meaningful insights from large-scale microscopy datasets. Researchers recognized the need for efficient data management, analysis, and interpretation of the increasing volumes of data. The field witnessed the emergence of interdisciplinary collaborations among biologists and computer scientists, leading to the development of specialized tools and platforms for image analysis, data management and visualization. In this regard, image databases, standardized file formats, and open-source software repositories became essential resources, fostering a collaborative and open scientific ecosystem. Accordingly, the advent of high-throughput microscopy and high-content screening propelled the demand for automated approaches. High-content screening emerged as a powerful technique for large-scale biological assays, allowing researchers to screen thousands of compounds and identify novel biological targets. Furthermore, the rapid advances in robotics, microscopy automation and computational resources enabled the analysis of enormous datasets having thousands or even millions of images. All of the above gave rise to powerful algorithms for object recognition and data mining, automatically developed to extract complex information, such as cell morphology, subcellular localization, and protein-protein interactions, at an unprecedented scale.

In the last decades, the integration of convolutional neural networks (CNNs), has revolutionized bioimage analysis. The ability of CNNs to automatically learn hierarchical representations from large amounts of labeled data has significantly surpassed traditional approaches for image segmentation, object detection and image classification [5]. The availability of large annotated datasets, combined with the advancements in computational hardware, has fueled the rapid progress of deep learning-based methods, opening up unprecedented possibilities for studying complex biological phenomena and accelerating the pace of discovery [6]. All of the above elucidates a bioimage analysis field which will undoubtedly witness further breakthroughs, driven by technological advancements, interdisciplinary collaborations, and the ever-increasing demand for more sophisticated analytical tools in microscopy.

1.1 Research Problem

Back in 2019 when this thesis began, the CNB (National Centre for Biotechnolgy) (http://www.cnb.csic.es/) held the distinction of being the largest research center within the Spanish National Research Council (CSIC), boasting considerable personnel and funding resources. By the end of 2016, the CNB had a workforce of 640 employees. Among its key strengths was the scientific-technical services platform, which offered cutting-edge technology in structural and cellular biology, genetically modified organism models, genomics, proteomics, computational biology, and synthetic biology. Over the past decade years, CNB researchers made impressive contributions, publishing over 1000 scientific articles, with 79% of them in top quartile journals within their respective fields. More than half of these publications were the result of international collaborations. The center boasts one member of European Molecular Biology Organization (EMBO) and eight European Research Counci (ERC) project leaders among its scientists. Recognized for its excellence in life sciences, the CNB ranked among the top three Spanish centers in the field according to the Nature Index and Scimago Institutions Ranking. Furthermore, the center was awarded the prestigious Severo Ochoa Center of Excellence accreditation in both 2014 and 2018.

The CNB possesses cutting-edge microscopy facilities at national and international level. In 2016, significant investments were made for a unique $200kV$ FEI Talos Arctica electron microscope, the sole one of its kind in Spain. This electron microscopy service provides European-level access through infrastructures such as iNext (https://inext-discovery.eu/), Instruct (https://instruct-eric.org/) and Corbel. A big amount of money was invested in the Advanced Light Microscopy Facility (ALMF) (https://www.cnb.csic.es/index.php/es/investigacion/servicios-cientificos/light-microscopy), which includes high-resolution microscopes such as total internal reflection fluorescence (TIRF) and stimulated emission depletion (STED) for gene expression detection and protein localization. These advanced microscopes can generate up to 1 terabyte of images per day depending on the operational mode. At that time, ALMF was one of the few facilities allowing external researchers to access these microscopes through research networks such as Campus Internacional de Excelencia UAM+CSIC, RedLab Madrid, and REMoA. Apart from these high-performance instruments, the AKMF service includes a Leica TCS SPs multispectral confocal microscope and a Leica Dl/l160008 fluorescence microscope, both equipped with an incubation system for experiments involving temporal changes.

In that context of growing need to process vast amounts of microscopy data, the Quantitative Image Analysis Unit for Microscopy (QIAU) (http://www.cnb.csic.es/index.php/es/component/k2/item/1669-quantitative-image-analysis-unit) was established at the end of 2019 with a clear purpose: to provide advanced image analysis and support, enhancing the capabilities of state-of-the-art microscopy equipment at both regional and national levels. The unit's foundation lied in recognizing the significance of quantifying processes and observed events through microscopy, as it plays a crucial role in advancing quantitative biology and objectively measuring hypotheses in cellular and molecular biology research. The proposed QIAU primarily focused on processing images acquired by the ALMF, which operated at an occupancy level of over 80% of the annual available microscope time. Furthermore, since 2009, the CNB had been an European reference center for image processing in the field of structural biology (Instruct Image Processing Centre (I2PC)). The software produced by I2PC (`http://scipion.i2pc.es`) has over 1000 users distributed worldwide. Dr. Sozano, who would oversee that new image analysis unit for light microscopy, was the technical director of I2PC. Thus the goal was to expand this strategy to light microscopy, by facilitating the transition from a predominantly qualitative and manual data analysis to a massive and quantitative analysis, fostering rapid advancements in Systems Biology. The aim was to complement the visual information provided by the acquisitions with quantitative parameters which enable objective access to the maximum numeric information. The technological challenges went beyond processing speed and automation of advanced image analysis operations, such as capturing the spatio-temporal complexity of cellular and subcellular dynamics. This included tasks such as particle and interaction detection, tracking, cell event detection and quantification of diffusion phenomena, and more.

1.2 Motivation

Until 2022, CNB-CSIC held the prestigious status of a Severo Ochoa Center of Excellence. This thesis is expected to make a significant contribution to the field of bioimage processing to become a leading center in quantitative biology and bioimaging. To achieve this, the center leveraged recently acquired cutting-edge microscopy equipment, encompassing both electron and optical microscopy, at the regional and national levels. Already recognized internationally for developing image processing algorithms for electron microscopy, the goal of this thesis was to extend automation and quantification capabilities to optical microscopy as well.

1.3 Objectives

The general goal of this dissertation, structured as a compendium of papers, emphasizes automating image processing techniques for fluorescence microscopy images acquired at the Advanced Light Microscopy Unit. The primary goal was to enable the transition from a predominantly qualitative and manual analysis to a massive and quantitative analysis approach. The secondary goal of the thesis was to identify the limits of current state-of-the-art of bioimage analysis and potentially extend these techniques to overcome their constraints. This was achieved by developing algorithms to address deficiencies observed in routine quantitative analysis of fluorescence images at ALMF. In particular, the following research topics have been identified:

1. *Extend Automation Capability*. One of the primary objectives is to significantly enhance the automation in image processing algorithms, focusing on tasks such as particle analysis, event counting, co-localization, spatial statistics, single-particle tracking, relative fluorescence quantification and cell-type classification. These advancements are aimed at optimizing the efficiency and accuracy of analysis tasks in ALMF. By automating these processes, researchers can achieve more streamlined and reliable results, leading to more meaningful investigations.

2. *Transitioning to Massive and Quantitative Analysis*. The aim is to move from the conventional qualitative and manual analysis and embrace a quantitative and massive analysis for optical microscopy images. This involves establishing a new quantitative image analysis unit as the central focus, which plays a vital role in the CNB's plan to become a leading center in quantitative biology and bioimaging. By adopting a more objective and high-throughput approach to analysis, this transition promises to provide more robust and advanced analyses to explore complex biological systems.

3. *Enhance the capabilities of bioimage processing pipelines at the ALMF*. By integrating state-of-art advanced image processing techniques into the existing pipeline, the thesis aims to bridge the gap among cutting-edge image processing techniques and practical applications in biology. This will involve adapting, optimizing, and developing necessary algorithms to ensure compatibility and functioning within the existing framework. The thesis seeks to empower researchers with more robust tools to encompass a variety of imaging modalities and biological specimens, demonstrating the versatility.

4. *Implement Real-Time Processing of Fluorescence and Super-resolution Microscopy Images*. The ultimate goal was to achieve streaming processing of fluorescence and super-resolution images acquired in the ALMF. This goal entailed implementing real-time image registration techniques to compensate geometric deformations induced by chromatic aberration while the TIRF microscope is acquiring. By accomplishing this, it provides immediate aberration-corrected data while the images are being acquired by the microscope on-the-fly, similar to the existing real-time processing in electron microscopy at CNB.

1.4 Outline of the Thesis

This thesis is organized as follows:

- *Chapter 2* covers the fundamental principles of image formation in optical microscopy, with a particular focus on the application of fluorescence microscopy in biological research. The chapter presents key concepts and techniques used in optical microscopy. It delves into the basics of image formation, including the role of aberrations in optical systems, and introduces wavefront, Zernike polynomials, and the point spread function. Moreover, it covers the fundamentals of fluorescence excitation and emission, Stoke's Law, Jablonski diagrams, and challenges such as photobleaching in fluorescence microscopy. Additionally, the chapter explores fluorescent labeling techniques and various fluorescence microscopy methods.

- *Chapter 3* is dedicated to comprehending the principles of bioimage analysis. It starts by covering the fundamentals of computer vision, digital images as functions and explaining the formation process. The concepts of digital images in spatial and frequency domains are also discussed. Additionally, this chapter introduces the impact of deep learning and neural networks in the field of bioimage analysis. It then proceeds to outline the essential stages in a typical bioimage processing and analysis pipeline. Furthermore, this chapter explores automated solutions to tackle challenges associated with large and multi-dimensional image datasets, as well as real-time processing. Emphasize the significance of open-source software, and common open-source solutions utilized in bioimage analysis are presented.

- *Chapter 4* presents the novel image processing and bioimage analysis developed during this thesis.(***No entiendo lo de los 50 anos). First, we discuss our contribution to cell-type classification using the `Cell-TypeAnalyzer` plugin. Next, we explore our work on single-particle tracking and motion classification using `TrackAnalyzer` plugin. Lastly, we detail our contribution to real-time chromatic aberration compensation with the `OFM-Corrector` protocol.

- In *Chapter 5*, a structured compendium of articles and co-authored publications used for this dissertation is presented, along with a brief overview of each, listed in chronological order.

- In *Chapter 6*, the primary conclusions of this dissertation are presented, along with potential avenues for future research

Chapter 2

Optical Microscopy: Shedding Light on the Microscopic Realm

Biology relies heavily on observations of natural phenomena to form and validate models of biological processes. While early observations were done in real-time and in-place, the value of keeping a visual record became apparent, as it removes spatial and temporal constraints. This led to the development of current imaging techniques, which are continuously evolving since the 20^{th} century, and providing new insights into biological dynamics with unprecedented spatial and temporal resolution [7]. Biological imaging field has greatly evolved over time, allowing observation of objects at a huge range of wavelengths, 3D geometries and nanometer-scale structures. Thereunder, different forms of microscopy can be used to observe cells and their internal substructures, ranging from advanced light microscopy to 3D cryo-imaging of native frozen samples [8]. The use of Green Fluorescent Proteins (GFP) has opened up a whole new color palette to be used in fluorescence microscopy, enabling the study of protein dynamics in living systems by genetically tag protein components. This chapter serves as an introduction to the field of optical microscopy, offering a comprehensive overview of its principles, components, techniques and applications. We will delve into the fundamental concepts of light and its interaction with matter along with image formation fundamentals understanding how these principles form the basis of optical microscopy. We will explore the key components of an optical microscope, the various microscopy techniques which leverage the power of light, advantages and capabilities of each technique, thus opening up new possibilities for studying different types of specimens and phenomena. We will also touch upon emerging trends and advancements in the field which push the boundaries of optical microscopy and enable even finer details to be revealed.

2.1 Unlocking the Hidden World: An Overview of Optical Microscopy

Optical microscopy is a widely used technique for exploring the microworld in biology, enabling researchers to understand the intricate biological samples on a small scale. By utilizing visible light and optical components, it magnifies images, allowing precise observation and analysis of microscopic structures. Optical microscopy offers non-invasive visualization of living samples, visualizing specific cellular structures without disrupting delicate biological processes [9]. This is achieved through fluorescent labeling or brightfield illumination. Recent advancements in hardware, algorithms and innovative approaches have

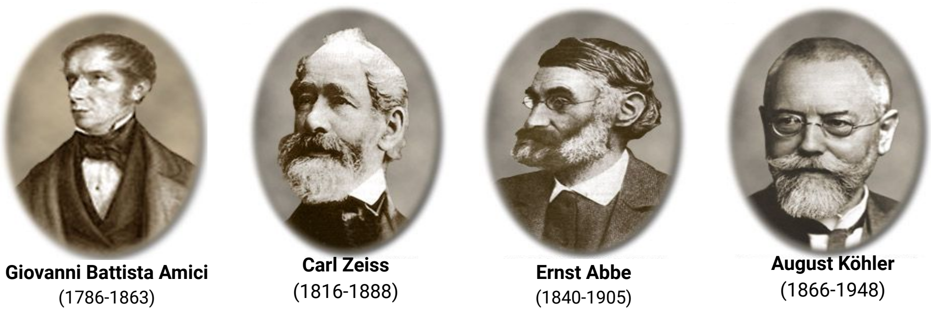

Fig. 2.1 Historical Perspective of Optical Microscopy: Giovanni Battista Amic, Carl Zeiss, Ernest Abbe and Agust Köhler

expanded the capabilities of optical microscopy, offering a wealth of information about cellular structures, dynamics and interactions in cell biology [10–12]. Early advancements in optical microscopy, such as achromatic objectives [13] by Lister and Amici (illustrated in Fig.2.1), and collaborations between pioneers such as Ernst Abbe (illustrated in Fig.2.1), Carl Zeiss (illustrated in Fig.2.1), and Professor August Köhler (illustrated in Fig.2.1), led to the development of apochromatic lenses and optimized photomicrography. The late 19^{th} century witnessed further innovations, such as metallographic microscopes, anastigmatic photolenses, binocular microscopes with prisms, or the first stereomicroscope [14]. In the early 20^{th}, parfocalized objectives were introduced, and Zeiss pioneered LeChatelier-style metallographs with infinity-corrected optics. Phase contrast microscopy gained recognition in the 1950s and remains popular in cell biology, enabling time-lapse cinematography of cell division.

2.2 Fundamentals of Image Formation in Optical Microscopy

In optical microscopy (Fig.2.2), image formation involves light interacting with the specimen and microscope components. Light from a source (e.g., lamp) converges onto the specimen, enhancing contrast. Interacting with the specimen's structure and composition, light is absorbed, transmitted or diffracted. According to Abbe's theory of image formation, diffracted light, out of phase (about 180 degrees) with direct light, leads to destructive interference at the image plane, creating light and dark patterns.

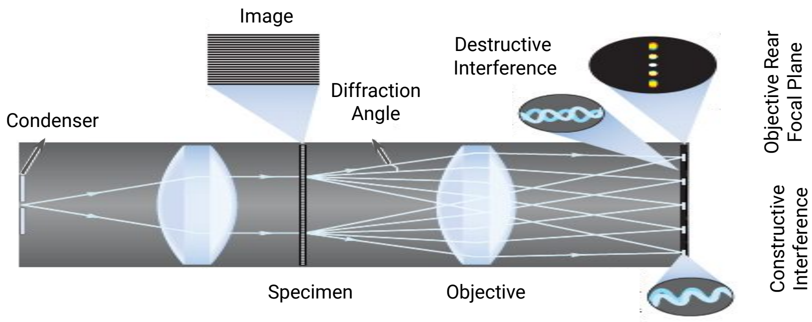

Fig. 2.2 Image Formation in an Optical Microscope

The objective lens is critical in optical microscopy, collecting transmitted or diffracted light from the specimen and focusing it to create an intermediate image, which is further magnified. Its numerical aperture (NA) determines resolving power and ability to capture fine details, as it indicates the light acceptance angle, affecting light gathering, resolution, and depth of field. The magnified image is projected onto an imaging plane (e.g., retina, camera film, or computer chip). The distribution of light and dark areas in the image reveals valuable information about the specimen structure and composition. Techniques such as staining or phase contrast imaging can enhance image contrast and reveal finer details [13]. Thus understanding image formation principles in optical microscopy unveils the complexity of the microscopic world.

2.2.1 Aberrations in Optical Systems

Optical systems typically designed with paraxial optics may overlook optical aberrations caused by light interacting with lenses [13]. Real optical systems deviate from this ideal path, presenting aberrations. These aberrations are due to: (1) the real path traveled by the light rays thought the optical system given by the exact application of Snell's law and (2) the refractive index variations as a function of light wavelength. Besides optical aberrations, other factors such as imperfections of microscope components, relative index mismatch, manufacturing defect, and environment factors degrade the optical performance, by impacting resolution, contrast and image quality in microscopes. Hence quantifying aberrations to be further compensated is crucial to enhance microscope performance. In this context, aberrations can be monochromatic or chromatic, stemming from lens or mirror geometry and occurring during reflection or refraction, even with chromatic light.

Geometrical Aberration: Five Seidel Aberrations In 1857, Seidel identified five constituent aberrations, known as the five Seidel aberrations, for first-order monochromatic aberrations [15, 16].

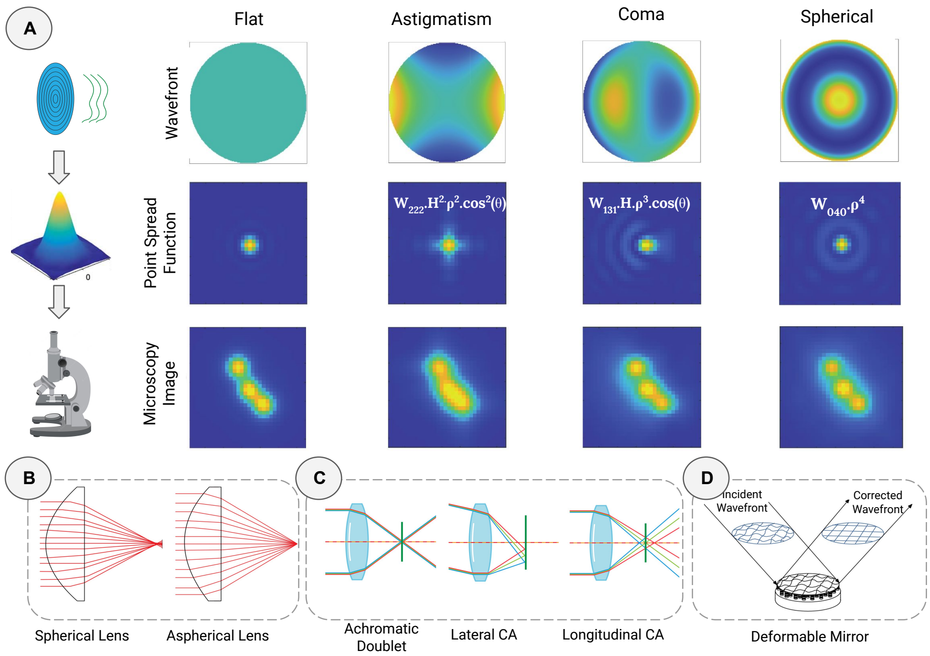

Fig. 2.3 Aberration in optical systems. (A) Examples of showing the effect of three common optical aberration modes together with corresponding wavefront coefficients. (B) Correction of Spherical Aberration by using Aspherical Lenses. (C) Correction of Lateral and Longitudinal Chromatic aberration by using achromatic doublet. (D) Deformable mirror.

- *Spherical Aberration.* It is a significant in objectives, resulting from the inability of a spherical lens to focus all incoming light to the same focal point on the optical axis [17]. It leads to a blurred image with reduced resolution and contrast, causing the specimen image to appear hazy and slightly out of focus (see Fig. 2.3 (B)). Correction methods include using aspheric lenses, counteracting with defocus, by lens splitting, or with higher index glass.

- *Coma Aberration.* It occurs due to refraction differences as light rays coming from an out-of-axis object point pass through different zones of a lens, particularly when magnification changes across the image and when the microscope is misaligned [18]. This distortion causes asymmetry in the image of an object point resulting in coma-like shapes (see Fig. 2.3 (B)). Correction methods include using a spaced doublet lens with a stop in the center.

- *Astigmatism.* It occurs when a thin bundle of rays, that strikes the lens surfaces obliquely, forms an astigmatic bean after refraction with two main focal lines: one in the sagital direction and one in the tangential direction. It is characterized by the off-axis image of a point appearing as a line or ellipse. So it appears elongated in one direction and compressed in the perpendicular direction (see Fig. 2.3 (B)) [18]. Correction methods include using a cylindrical lens, higher index glass, counteracting with defocus, or using a spaced doublet lens with a stop in the center.

- *Curvature of Field.* This aberration appears when lens elements focal lengths (multiplied by refractive indices) not summing up to zero. Modern microscopes use specially designed objectives, such as plano. The image of a set of extra-axial points may be perfect but formed on a non-planar surface (Petzval surface)[18]. It can be corrected with spaced doublet lenses.

- *Distortion.* Even after compensating all aberrations above, it is possible for light emanating from points within the object to converge on the image plane at an incorrect distance from the optical axis. This divergence from linear proportionality with the object's distance from the optical axis can result in two types of distortions. If the distance increases more rapidly than in the object, it leads to pincushion distortion, causing objects at the periphery to appear stretched or elongated towards the image edge. Conversely, if this increase is slower, it results in barrel distortion, causing objects near the center of the field of view (FOV) to appear compressed or squeezed.

Chromatic Aberration CA is a common optical aberration caused by lens dispersion affecting quality of images acquired with optical microscopes. It occurs when the lens focus shifts with the light wavelength due to chromatic dispersion, resulting in different wavelengths being focused at various positions due to their different speeds while passing through the optical lenses. As a result, different colors will come to focus at slightly different planes, leading to blurred and colored edges around objects. In this context, longitudinal CA (illustrated in Fig.2.3 (C)) occurs when light of different wavelengths does not converge at the same focal point along the optical axis.

> This means that the different colors appear at different distances from the microscope objective, resulting in color fringing along the axial direction of the image. This type of CA is more pronounced in high NA lenses and particularly problematic with high magnification objectives. Lateral CA (illustrated in Fig.2.3 (C)) occurs when light of different colors focuses at different lateral positions in the image plane. Unlike longitudinal CA, this aberration becomes more evident towards the periphery, as color fringing along the edges of the specimen or the FOV. Therefore, reducing CA is crucial for acquiring high-quality images with accurate color representation. Several techniques are employed to minimize CA such as achromatic and apochromatic lenses (illustrated in Fig.2.3 (C)) which combine multiple lens elements made from glasses with different chromatic dispersion to bring different colors of light to a common focus, reducing both longitudinal and lateral CA. Some modern microscopy systems and image processing software offer post-processing algorithms to correct CA. The use of monochromatic light sources completely eliminates CA since only a single wavelength is used, hence all rays converge to the same focus.

2.2.2 Wavefront and Zernike Polynomials

Optical aberrations result from deviations in the wavefront at the exit pupil of an optical system compared to the ideal wavefront of a perfect optical system [19]. When the ideal wavefront is spherical (non-aberrated), rays from the object point converge to the image point. However, if the wavefront deviates from spherical, the image becomes aberrated, and rays do not follow the same optical path, leading to image degradation. The wavefront, defined at the exit pupil, (described in Fig.2.4(A)) is used to mathematically model the image quality of an optical system, and the wave aberrations defined as the difference among the actual wavefront and the spherical one describe the aberrations of the optical system [19]. In optics, the wavefront of an optical system is described using a linear combination of Zernike polynomials [20] (Fig.2.4 (B)). Zernike polynomials, in polar coordinates $(x = \rho cos\theta, y = \rho sen\theta)$, are orthogonal functions representing wavefront shape in terms of coefficients. As most optical systems have rotational symmetry with circular pupils, these polynomials are suitable for characterization and correction of aberrations. The wavefront $W(\rho, \theta)$ in polar coordinates can be expressed as a polynomial expansion [21], as described in Fig.2.4 (C).

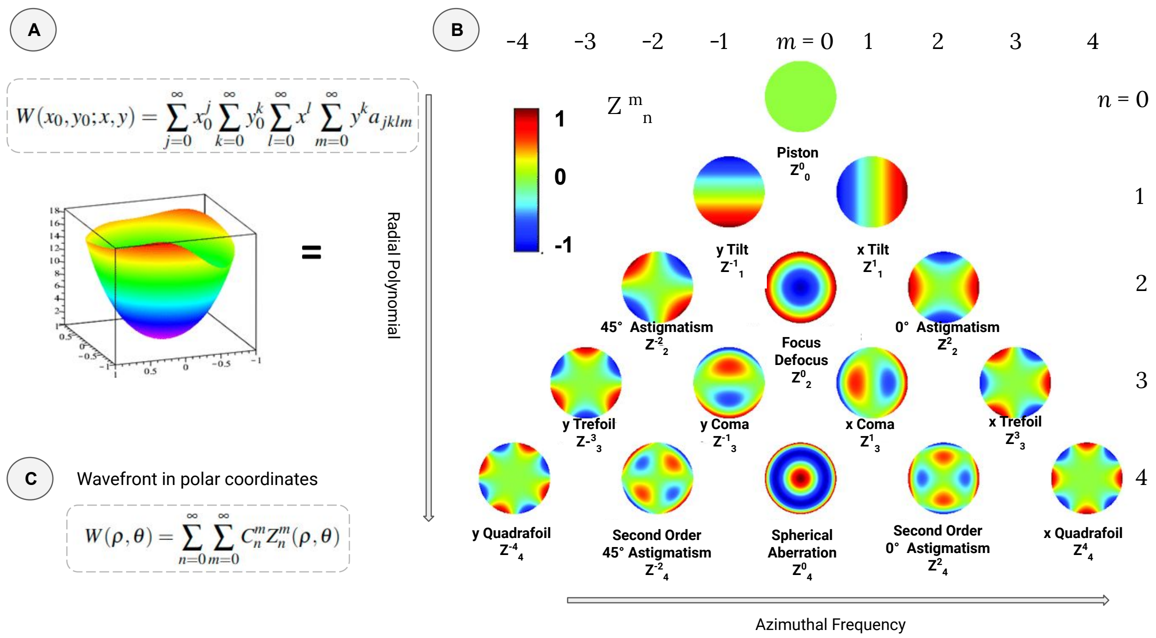

Fig. 2.4 Wavefront and Zernike Polynomials. (A) Wavefront of an optical system is described using Zernike polynomials. (B) The first five orders of aberrations as their Zernike Polynomials Z_n^m. The colorbar shows the value of the functions. (C) Wavefront $W(\rho,\theta)$ in polar coordinates can be expressed as a polynomial expansion.

2.2.3 Point Spread Function

When using optical microscopy, the acquired image is a blurred representation of the object due to inherent aberrations in the microscope. The point-spread-function (PSF) describes the appearance of a point of light emitted by the specimen when observed through the microscope [22]. The PSF is a 3D diffraction pattern of light (shown in Fig.3.1 (A-B)) emitted from a point source and transmitted to the image plane through the objective lens. In non-aberrated optical systems, the diffraction pattern is periodic and symmetrical in both the axial and lateral planes at the paraxial focal point [23]. However, the diffraction pattern can have various shapes depending on the imaging system. The axial and lateral resolutions can be assessed using the PSF (shown in Fig.3.1 (B)), which is generated from optical sections along the z-axis [23]. As optical microscopy follows a linear and shift-invariant image formation process, this property enables the image computation through a convolution process as follows:

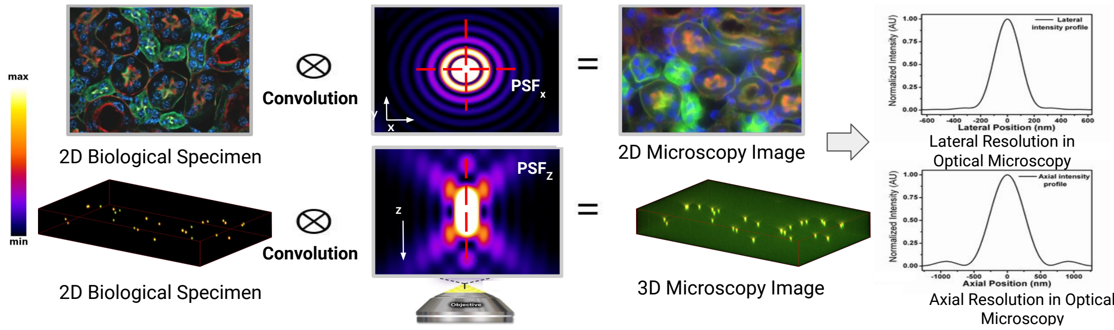

Fig. 2.5 Point spread function of an optical microscope. Lateral and axial perspectives. In the Cartesian reference system, the origin is positioned at the peak center of the PSF. The lateral view corresponds to the plane where z equals zero, while the axial view corresponds to the plane where y equals zero. Axial (with z and y equal to zero) and lateral (with x and y equal to zero) intensity profiles.

The image degradation can be modeled by assuming a perfect image f blurred by convolution with a kernel h and corrupted by noise ε:

$$\hat{f}(x,y) = f(x,y) \circledast h(x,y) + \varepsilon(x,y) \tag{2.1}$$

The convolution kernel h or PSF, models blurring caused by degradation sources (FOV and imaging device). Since the PSF is always normalized, it is straightforward to compare the PSF of different systems and assess their respective imaging capabilities. The PSF plays a crucial role in characterizing the microscope's resolution and imaging capabilities, being typically modeled with a Gaussian Function [24] or by measuring the full-width at half-maximum (FWHM) of the PSF which measures the distance among the points where the intensity is half of the maximum.

The Airy Disk and size of the PSF The Airy disk is a 2D diffraction pattern seen when a point source of light is imaged through an optical system. It has a central bright spot surrounded by concentric rings and sets the resolution limit. The size of the Airy disk is determined by the radius of the Airy disk's central maximum as $r_{airy} = \frac{0.61\lambda}{NA}$. While the Airy disk represents the diffraction limit and has a fixed size determined by the NA and wavelength, the PSF is a broader concept which characterizes how an optical system responds to a point source of light, whose size and shape can vary, being affected by optics quality or aberrations present in the system

2.3 Fluorescence Microscopy

Fluorescence microscopy is a dominant technique in life sciences, enabling visualization and study of specific molecules and structures within cells or tissues. Unlike brightfield microscopy, it utilizes fluorescent molecules or fluorophores, which absorb excitation light of a specific wavelength and re-emit it with a longer wavelength (phenomenon called fluorescence), allowing selective labeling and precise visualization with exceptional sensitivity. The process of imaging with fluorescence microscopy involves several steps. First, introducing into the sample fluorescent probes which are specific to the target of interest. They can be designed to bind to particular molecules (e.g., antibodies binding to specific proteins) or to target specific cellular structures. Then fluorophores are excited with specific-wavelength light, and the emitted fluorescent light is collected by the objective lens. To ensure that only the emitted fluorescence is observed, filters are used to selectively transmit the emitted light while blocking the excitation light. Finally, the emitted light is magnified, and detected with a camera or photomultiplier to convert the light signal into an electronic image.

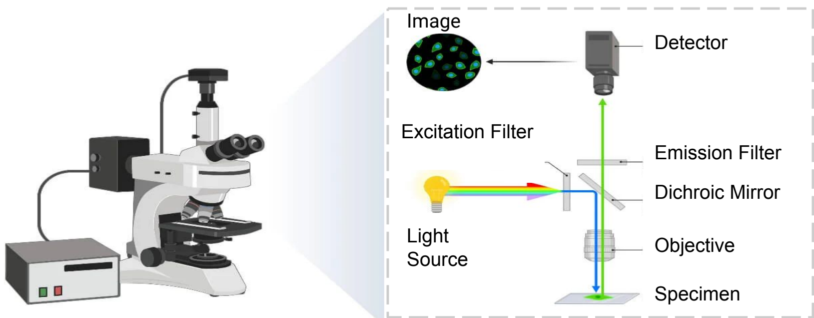

Fig. 2.6 Process of imaging with fluorescence microscopy. Image is formed by focusing the emitted flourescence light into a detector which is an electronic system.

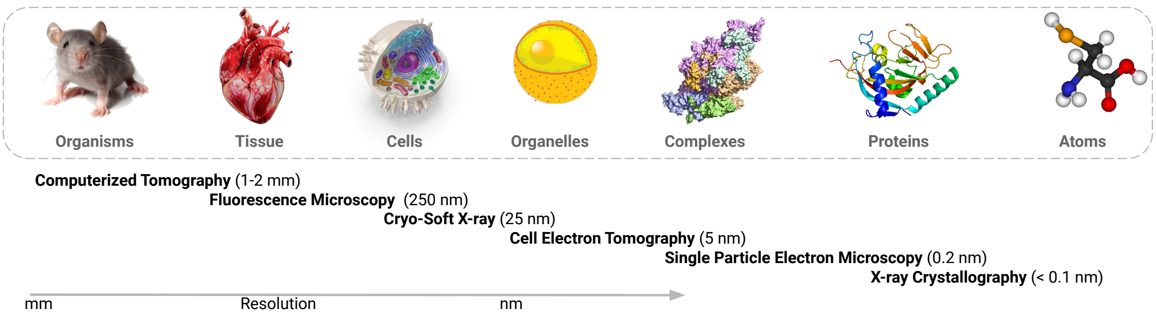

Fig. 2.7 Biological imaging involves a range of techniques developed to explore the structures within biological systems across various scales, organized by the achievable levels of resolution.

Fluorescence microscopy offers advantages over other techniques such as exceptional sensitivity to detect low concentrations of labeled molecules within complex samples. It also enables real-time visualization of dynamic processes [25] and finds widespread applications, as the understanding of cellular structures, molecule localization, interactions and biological process dynamics [26]. Advances such as confocal and super-resolution microscopy have further enhanced resolution, making this technique core in many microscopy facilities. In the following lines, we will explore the principles, instrumentation and advanced techniques associated with fluorescence microscopy. We will delve into the various labeling strategies, imaging modalities, shedding light on its immense potential for scientific discovery.

2.3.1 Fluorescence Excitation and Emission Fundamentals

Fluorochromes are chemical compounds having photoreactive properties, absorbing light at a particular wavelength and emitting light at a longer wavelength. This make them precious as detection agents. Fluorochromes have distinct absorption and emission spectra (usually similar to excitation) due to their electronic configurations. Manufacturers specify peak excitation and emission wavelengths for each fluorochrome.

Emission Spectrum of a Fluorochrome

To analyze the emission spectrum of a specific fluorochrome, it is needed to identify firstly the wavelength at which it exhibits maximum absorption, usually corresponding with the excitation peak. The fluorochrome is then stimulated at this wavelength to initiate excitation. In Fig.2.8, we can observe the absorption spectrum of a typical fluorochrome. The excitation spectrum of the fluorochrome is determined by monitoring the fluorescence emission at the wavelength of maximum intensity while exciting the fluorophore with a sequence of consecutive wavelengths. The emission maximum is selected, allowing only the emission light at that particular wavelength to reach the detector. The intensity of emitted fluorescence is quantified by exciting it at different excitation wavelengths, and then recording it as a function of wavelength. The outcome is a curve, illustrated in Fig.2.8, illustrating the relative fluorescence intensity resulting from excitation across the spectrum of excitation wavelengths.

2.3.2 Stoke's Law or Shift

When electrons transition from an excited state (S_1, S_2) to a ground state (S_0), vibrational energy is lost. This loss of energy causes the emission spectrum to shift towards longer

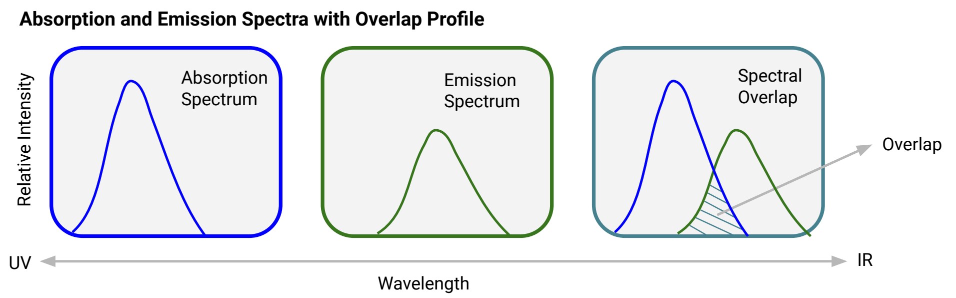

Fig. 2.8 The absorption, emission and excitation-emission overlap spectrum of a typical fluorochrome.

wavelengths compared to the excitation spectrum, as wavelength is inversely proportional to radiation energy according to $E = \frac{hc}{\lambda}$. Stokes' Law, or the Stokes' shift, causes the emission spectrum to shift towards longer wavelengths due to energy loss during electron transitions. A larger Stokes' shift helps separate excitation from emission light. Fluorescence intensity is maximized by exciting the fluorochrome at its peak excitation wavelength and detecting emitted light at the peak emission wavelength (or other selected wavelengths). Filters are used to regulate excitation and emission wavelengths. Figure 2.9 shows similar-shaped fluorescence intensity curves for absorption and emission, with overlapping excitation and emission curves at specific wavelengths.

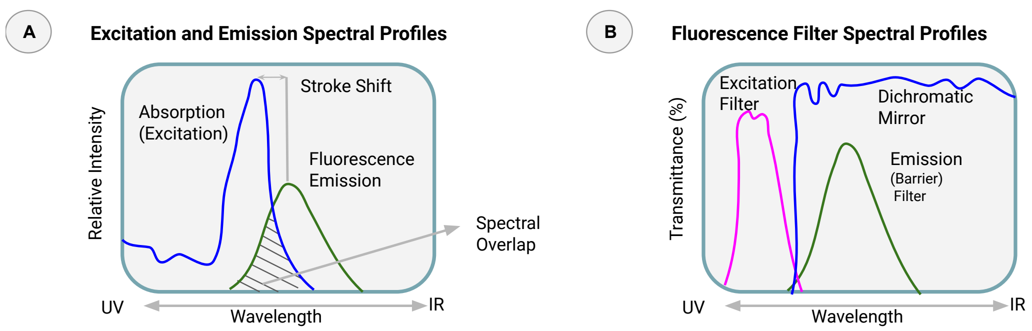

Fig. 2.9 Observations from Excitation and Emission Spectrums. (A) A typical fluorochrome absorption-emission spectral diagram. (B) Fluorescence Filter Spectral Profiles.

2.3.3 Separation of Excitation and Emission Wavelengths

Proper filter selection allows the separation of excitation and emission wavelengths (Fig.2.9). Fluorescence illuminators use interchangeable filters into the light path to control light before reaching the specimen (excitation) and as it emanates from the specimen (emission). Using a bright light source for excitation and fluorochromes with satisfactory absorption as well as yield, maximizes weak emission light. The molecular extinction coefficient determines the efficiency of a fluorochrome absorption of excitation light, crucial for subsequent fluorescence emission. The quantum yield represents the ratio of emitted quanta (energy packets) to absorbed quanta for the emitted light [27].

2.3.4 Molecular Explanation of Fluorescence: Jablonski Diagrams

The Jablonski energy diagrams [28] (described in Fig.2.10(B)) explain the physical relation and energy transitions between light absorption and emission from a fluorophore, showing the different energy levels involved in the photons' absorption and emission. Representing energy levels with a vertical axis and transitions with horizontal arrows, these diagrams provide insights into the photophysical properties of fluorophores [29].

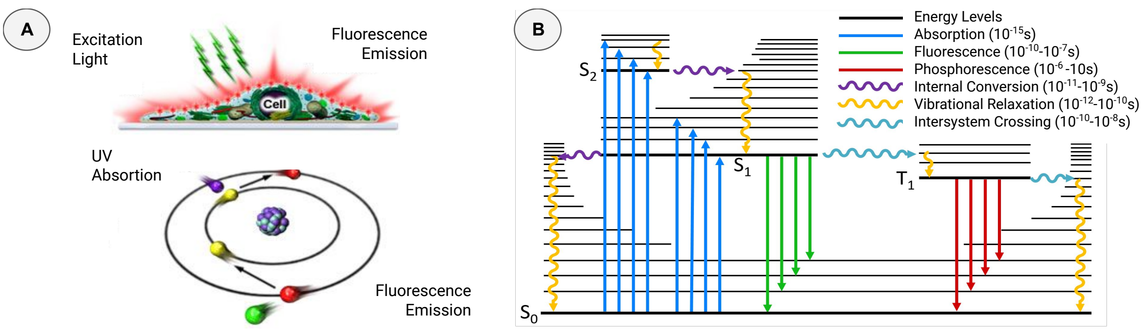

Fig. 2.10 Fundamentals concepts underpinning Fluorescence Microscopy. (A) Cartoon of fluorescence technique and Stokes Shift. (B) Examplo of Jablonski diagram showing the possible radiative and non-radiative transitions.

Fluorophores, normally in the ground state S_0 (absense of excitation), get excited to higher energy levels S_1 and S_2 after absorbing a photon, typically within a few femtoseconds (10^{-15} seconds),represented by an upward arrow in the diagram. Following excitation, rapid non-radiative transitions such as vibrational relaxation occur, where it loses excess energy through molecular vibrations. This leads to fluorescence emission as the fluorophore returns to the ground state (S_1 level from S_0 within *picoseconds*), emitting a photon of lower energy (with longer wavelength) than the absorbed photon [30, 31]. In some cases, the fluorophore may undergo internal conversion, a non-radiative transition from the excited state S_1 back to the ground state S_0. However, in most cases, this transition among excited and ground states results in a fluorescence emission. In this case, after the fluorophore has been excited, it returns to the ground state S_0 by emitting a photon of lower energy (with longer wavelength) than the absorbed photon. This emission is represented by a downward arrow in the diagram. In this regard, the time the fluorophore stays in the excited state before emitting a photon is called the fluorescence lifetime. This lifetime is directly influenced by factors such as the environment, molecular interactions and the presence of quenchers.

2.3.5 Dealing with Fading or Photobleaching

Photobleaching is the irreversible photochemical damage of a fluorophore due to light intensity and molecular oxygen, causing a permanent fading of fluorescent signal [29]. This leads to two types of artifacts at the molecular level: 1) Fluorophores suddenly disappearing, resulting in an illusion of faster diffusion and reduced residence times in the detection volume [32, 33]. 2) Gradual depletion of fluorescence within enclosed small volumes such as cells or vesicles, and even in two-dimensional systems (membranes), where fresh molecules' diffusion cannot compensate for fluorophore depletion, leading to distortion in correlation curves [33]. While photobleaching is inevitable, various approaches can be employed to mitigate its effects in imaging certain specimens.

- *Reducing the light intensity.* To reduce light intensity during imaging until only a portion of fluorophores is bleached. This minimizes photobleaching by reducing excitation-emission cycles. However, a balance is needed, as lower excitation light also means lower signal intensity and contrast.
- *Reducing the exposure time.* By reducing the exposure time to light as it will decrease the times the fluorophores undergo the excitation-emission cycles. This can be achieved either by decreasing pixel dwell of laser or by choosing a faster imaging frame rate.
- *Adding Anti-Fade Reagents.* Immersing a sample on a mounting medium may reduce photobleaching. Yet, not all dyes respond equally to anti-fade reagents, so the choice of anti-fade agent should be tailored to the specific dye being used.
- *By using Neutral Density Filters.* These filters are used in the light path before the light reaches the excitation filter to decrease excitation intensity. This efficiently allows the passage of almost all emitted wavelengths by reducing photobleaching. However, this may also decrease the sample's signal while reducing its exposure to light.

2.3.6 Fluorescent Labeling Methods

Fluorescent labels are essential for fluorescence-based assays, enabling selective detection, visualization, and monitoring of non-fluorescent cell types, dynamic cellular processes, or subcellular structures. While proteins present intrinsic fluorescence (typtophan), often extrinsic fluorescent labels are needed as imaging agents to enhance fluorescence properties during imaging [31, 34]. Various fluorophore labeling techniques exist, as many as protein diversity, each with its advantages and considerations based on the microscopy technique and biological system used [35].

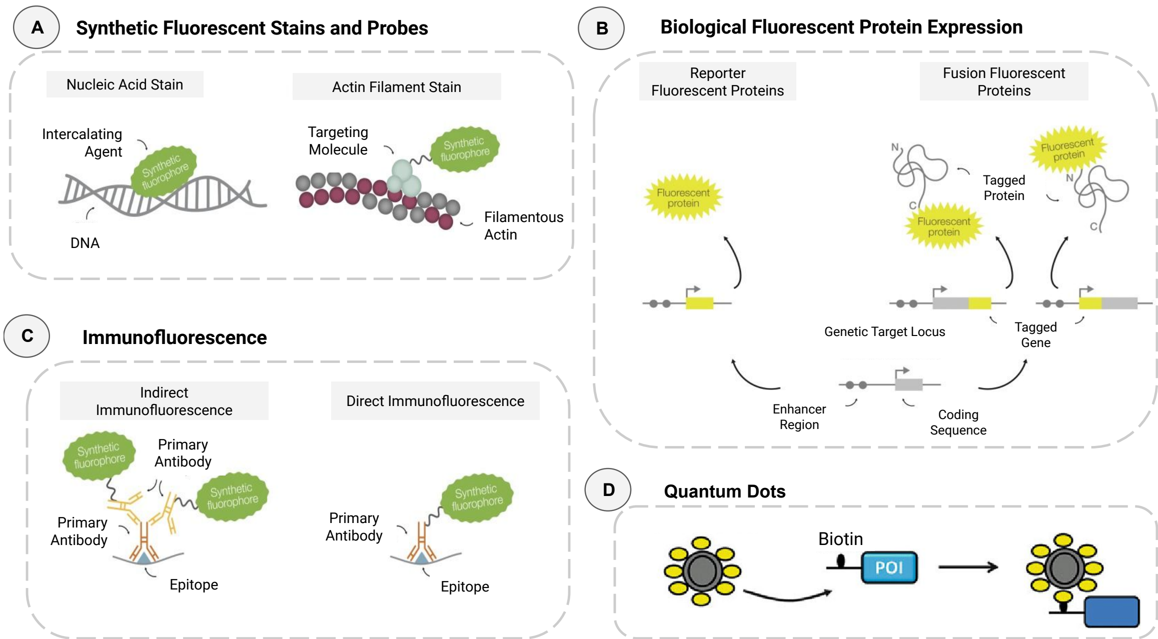

Fig. 2.11 Fluorescent Labeling Methods. (A) Fluorescent stains and probes interact with cell components or can merge with targeting molecules. (B) Immunofluorescence pairs synthetic fluorophores with immunoglobulins that selectively bind to specific target protein antigens. They can also attach to a primary antibody for direct interaction (direct immunofluorescence). (C) Fluorescent proteins expressed naturally offer a genetic way to visualize cell components. (D) An example: biotinylation protein conjugation with streptavidin-functionalized quantum dot.

- *Biological Fluorescent Proteins.* Derived from biological structures (Fig. 2.11(B)) can be attached to target proteins (e.g., proteins, enzymes, or antibodies) for labelling [36]. They can be bound into in-vivo proteins and introduced within living cells, bacteria, or organisms, with lower toxicity compared to synthetic dyes. Recent developments include photoactivatable, photoswitchable, and photoconvertible fluorochromes for studying protein dynamics and for single-molecule based Super-resolution microscopy [30].

- *Synthetic Fluorescent Stains and Probes.* Synthetic dyes and probes (Fig. 2.11(A)) are widely used for imaging fixed cells or tissues, selectively staining nucleic acids, lipids, cellular structures and organelles. For live imaging, cell permeability of the stain or probe is crucial. Commonly used examples include Hoechst 33258 and DAPI for nucleic acids, NileRed and FM dyes for biological membranes, fluorophore-derivatized phallotoxins for actin filaments and LysoTracker, MitoTracker, and ER-Tracker for organelle labeling [30].

- *Immunofluorescence Staining.* Antibody-based staining (Fig. 2.11(C)) is a commonly used technique, employing synthetic fluorescent dyes coupled to immunoglobulins to visualize proteins in fixed cells or tissues. It offers higher specimen contrast and signal amplification, allowing flexibility in choosing fluorescent dyes with different wavelengths. However, this method is unsuitable for live imaging due to the need for prior fixation and membrane permeabilization. Multicolor immunofluorescence enables simultaneous visualization of multiple cellular components [30].

- *Graphene Quantum Dots.* GQD (Fig. 2.11(D)) are fluorescent nanomaterial with unique optical and electronic properties for imaging and sensing applications [34]. Despite being stable, compared to organic dyes [37], GQD is a new procedure which requires improvement for toxicity due to their heavy metal composition and high stability.

2.3.7 Fluorescence Microscopy Techniques

Throughout the following lines of this section together with the Fig.2.12, it will be found a brief and technical description of the commonly used fluorescence microscopy techniques.

Widefield Microscopy. This technique (Fig. 2.12(A)) is widely used for studying large-scale biodynamics [38] in fixed or live cells, tissues, and organisms. It illuminates the entire FOV and collects emitted fluorescence on a detector. This method provides fast imaging and simple setup, with reduced photobleaching and phototoxic effects due to a small light dose for illumination [39, 30]. However, it also captures out-of-focus image information, compromising image resolution. To address this, structured illumination or post-acquisition methods such as deconvolution can be applied [40].

Optical Sectioning Microscopy. This methodology (Fig. 2.12(B)) revolutionized optical imaging by eliminating out-of-focus background light, resulting in improved resolution beyond widefield microscopy [41]. However, understanding optical and fluorescence concepts is crucial for obtaining high-quality images. It is comprised of the following techniques.

- *Confocal Laser Scanning Microscopy (CLSM).* CLSM uses a scanning laser beam with a pinhole aperture and photodetectors to reject out-of-focus signals. It consists of an epifluorescence microscope with laser sources, fluorescent filter sets, a scanning mechanism, pinhole apertures, and PMT detectors. The laser beam is raster-scanned on the specimen, and the emitted light passes through the pinhole aperture while out-of-focus light is rejected. The PMT converts the intensity into a digital signal for display. The lateral resolution is around 200 nm, depending on NA and illumination wavelength [40].

- *Spinning-Disk Confocal Microscopy (SDCM).* SDCM combines the out-of-focus light rejection of CLSM with the high sensitivity of widefield microscopy [42]. It is ideal for high-resolution imaging of small cells, significantly improving image contrast and signal-to-noise ratio (SNR) by eliminating out-of-focus light. SDCM has advantages over CLSM, especially for fast in vivo imaging and conditions requiring higher frame rates. Modern SDCM employs low-intensity excitation light and fast imaging to reduce photobleaching and phototoxicity simultaneously [43].
- *Multiphoton Microscopy (MPM).* This method encompasses the simultaneous absorption of two or more photons produced by NIR femtosecond pulsed laser excitation by a single fluorophore [44], producing high-resolution 3D images [45]. By eliminating light coming from out-of-focus planes and not requiring a pinhole near the detector, it requires small amounts of photons to illuminate the specimen. This efficiently reduces fluorophore bleaching and phototoxicity, enabling label-free real-time imaging of biological processes without cell damage [30].
- *Ligth-Sheet Fluorescence Microscopy (LSFM).* This microscopy technique limits photodamage in live-cell imaging by using a thin sheet of laser light to excite fluorophores only within a narrow plane, a few hundred nanometers to micrometers [46]. Fluorescent photons emitted by the fluorophores are captured by a detection objective positioned perpendicularly to the light sheet and then imaged onto a CCD. This fast acquisition provides high temporal resolution, minimal photobleaching, and reduced phototoxicity by illuminating only the observed plane [47, 39].

Super-Resolution Optical Microscopy. SRM methods (Fig. 2.12(C)) use advanced fluorescence imaging techniques to resolve objects beyond the diffraction limit [48]. They employ engineered excitation light, fluorescent dyes, sensitive detectors, faster processing, and reconstruction algorithms to reduce the size of the PSF [49]. Recognized with a Nobel prize in 2014 [50], these methods have enabled molecular resolution at the nanometer scale, fast live-cell imaging, and volumetric 3D multi-color imaging.

- *Super-Resolution Structured Illumination Microscopy.* SR-SIM [51] is a super-resolution technique which surpasses the diffraction limit [52]. Its strengths include compatibility with conventional sample preparation, efficient utilization of available photons with highly sensitive cameras, and reduced excitation power while achieving high-quality fluorescence detection. It extends widefield capabilities, allowing multi-color imaging (up to four color channels), optical sectioning, and fast live cell imaging with doubled lateral and axial resolution compared to optical sectioning microscopes [51].

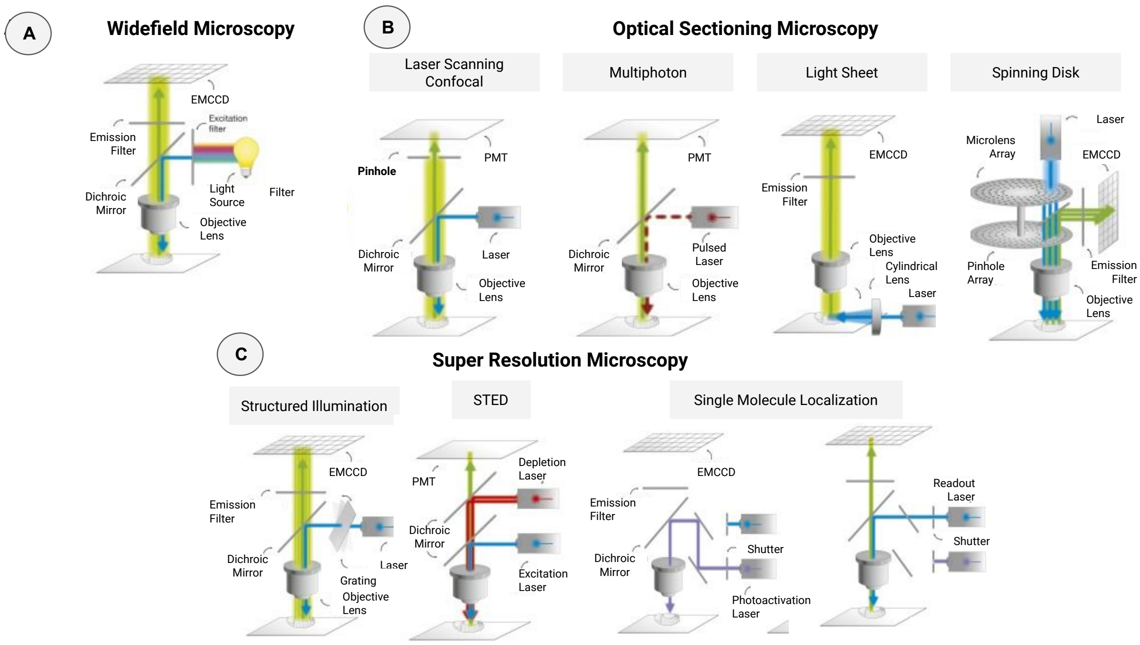

Fig. 2.12 Commonly used Fluorescent Microscopy Techniques. Fluorescence imaging technologies can be classified into three main categories: (A) Widefield Microscopy; (B) Optical Sectioning Microscopy; (C) Super-resolution Microscopy

- *Stimulated Emission Depletion (STED).* This imaging approach achieves sub-diffraction resolution below the limit of $\lambda/2NA$ [53], preserving optical sectioning and molecular specificity/sensitivity. It has made significant strides, considering the fluorophore as an active element in image formation [54]. However, photobleaching and phototoxicity at high intensities required need to be addressed, with a focus on minimizing excitation intensity and optimizing fluorophore properties for long-term and live-cell imaging. Improving temporal resolution and range is essential for capturing fast dynamic processes in living cells. STED faces limitations in thicker samples due to light scattering and absorption, necessitating advancements in optics and labeling strategies to penetrate deeper into tissue for broader biological applications.

- *Photo-Activated Localization (PALM) and Stochastic Optical Reconstruction (STORM) Microscopy.* PALM and STORM [55–58] are single-molecule localization-based super-resolution microscopy techniques. They use photoswitchable fluorescent dyes or proteins in presence of suitable buffers to achieve high-resolution imaging [59]. These methods stochastically activate a subset of fluorophores with a low-power laser, followed by photobleaching (PALM) or switching them into a reversible dark off-state (STORM) using an inactivation laser. Detected emission events are then used to reconstruct the final image. Both techniques achieve lateral resolutions of 20 nm and axial resolutions of $50 - 60$ nm. While PALM was initially developed with photoactivable or photoconvertible fluorescent proteins and STORM with synthetic dyes, both types are now interchangeable in both methods. PALM and STORM have been extended to multi-color and 3D imaging, making them highly applicable for a wide range of biological applications. However, careful probe selection is necessary, especially for multi-color imaging modes [30].

Chapter 3

Bioimage Analysis: Unveiling Insights from Biological Images

The concept of bioimage analysis emerged from the convergence of microscopy, computer science and quantitative biology. It employs computational methods to extract quantitative data from images of biological samples. The transition from analog to digital imaging in the late 20^{th} century, and benefiting from computer vision and image processing techniques, facilitates digital images to be stored and processed using computers. Researchers started applying image processing algorithms to extract meaningful information from biological images. The advent of high-throughput imaging technologies, such as confocal microscopy, marked a pivotal moment by enabling the acquisition of vast amounts of image data. This increased the need for automated and efficient bioimage analysis tools. Therefore, the concept of bioimage analysis emerged as a result of interdisciplinary collaboration between biologists, physicists, mathematicians, and computer scientists, to develop tools which could address the specific challenges of analyzing these complex biological images. Furthermore, in recent years, open-source software platforms, such as ImageJ and CellProfiler, democratized access to bioimage analysis tools. Additionally, the shift toward quantitative biology, with an emphasis on data-driven further propelled the growth of bioimage analysis. This field continues to expand, offering researchers a robust means to uncover insights into biological processes, aided by ongoing advancements in imaging technologies and computational methodologies.

This chapter explores how bioimage analysis, combining microscopy, computer science, and quantitative biology, revolutionizes insights from biological images. It covers the role of computer vision and it describes classical and cutting edge methods and applications to achive the results exhibited in this study. Moreover, open-source software's impact and its connection to quantitative biology are discussed.

3.1 Fundamentals of Computer Vision in Light Microscopy

Human beings possess the innate ability to perceive the three-dimensional world. With the advent of digital image capturing devices and high performance computers, it is nowadays possible to obtain information from images in a similar way as the human visual system does. Computer vision is a field belonging computer science which enables machines to interpret visual information from their environment, mimicking human visual perception [60, 61]. It encompasses topics such as image formation, feature detection, segmentation, object recognition, and tracking. Deep learning techniques such as Convolutional Neural Networks (CNN) have been widely adopted in contemporary computer vision applications, enhancing capabilities for wide range of image processing tasks.

3.1.1 Digital Image Formation

A digital image is a function which maps spatial coordinates to intensity values for grayscale or pseudo-color images. In the continuous domain, it is represented as $f(x,y) : \mathbb{R}^2 \to \mathbb{R}$, with (x,y) as coordinates and $f(x,y)$ as intensity/color index. In computer graphics, images are in the discrete domain as $F[m,n] : \mathbb{Z}^2 \to \mathbb{R}$, with (m,n) as pixel coordinates and $F[m,n]$ as intensity/color index. True color images are maps from $\mathbb{R}^2 \to \mathbb{R}^3$. For each position (x,y), the function $f(x,y)$ is a vector (r,g,b) with the coordinates of the color of the pixel in the *RGB* system. Unless specified otherwise, our images are grayscale.

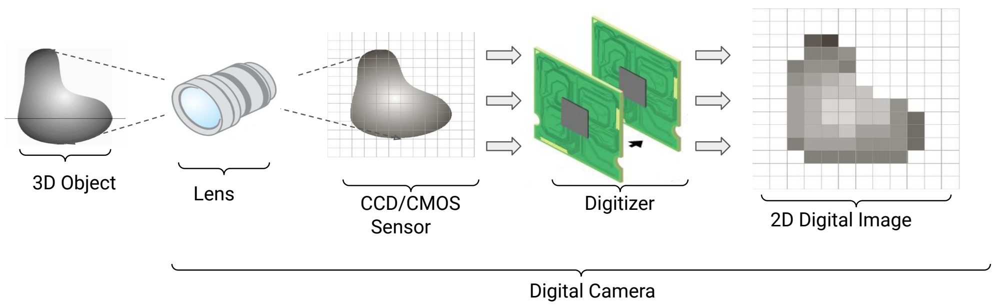

Fig. 3.1 Digital Image Formation. An image can be conceived as a 2D function $f(x,y)$, being x and y the spatial coordinates, and the amplitude of f at any pair of coordinates (x,y) is the image intensity at that level.

Digital image formation, as shown in Fig.3.1, involves two main processes: sampling and quantization, which convert the continuous analog image into a discrete digital image. Sampling captures the analog image at discrete intervals using an image sensor with millions of pixels. Whereas quantization assigns numerical values to pixels based on intensity or color, represented by binary codes (0 to 255 in an 8-bit image). The higher the bit depth, the more colors and shades of gray it can represent [60].

The mathematical model for digital image formation [62, 63] is given by:

$$f(x,y) = s(s,y) * h(x,y) + \varepsilon(x,y) \tag{3.1}$$

where $f(x,y)$ is the observed digital image, $s(x,y)$ is the continuous analog signal, $h(x,y)$ is the *PSF* that models the blur and distortion introduced by the imaging system, and $\varepsilon(x,y)$ is the additive noise introduced by the imaging device.

3.1.2 Digital Image Sensing

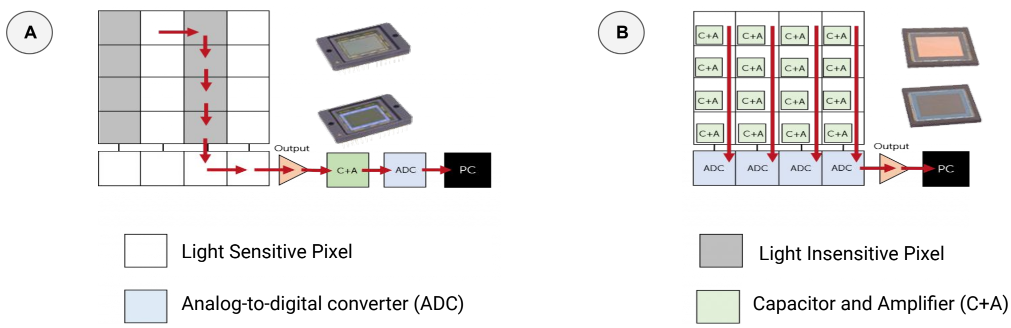

Fig. 3.2 Two common types of image sensors used in digital imaging are (A) CCD and (B) CMOS sensors

Digital image sensors capture light and convert it into electrical signals. These components consist of an array of photosensitive elements, representing incoming light at each pixel to be further processed, stored and displayed as images. Charge-coupled device (CCD) is a digital image sensor, which work by converting photons into electrical charge. Each pixel on a CCD sensor is a light-sensitive photodiode accumulating charge proportional to the light it receives to an output node. CCD are known for their excellent image quality, low noise, and suitability for applications with low-light conditions. Likewise, in Complementary Metal-Oxide-Semiconductor (CMOS), each pixel has its own amplifier, which amplifies the charge generated by the photodiode [63, 60]. This allows for parallel processing of pixel information, making CMOS faster than CCD, presenting lower power consumption. Additional components in this process include the lens system, the analog front-end, and the analog-to-digital converter (ADC). The lens system determines image quality, while the analog front-end amplifies and filters the analog voltage before ADC conversion.

3.1.3 Digital Image in Spatial Domain

A digital image is represented in the spatial domain as a 2D function of space, where pixel values indicate the amplitude at each pixel. Manipulation operations such as smoothing, sharpening, noise reduction, and edge detection (Fig.3.3) can directly manipulate pixel values in their spatial coordinates. Moreover, various image properties, such as histogram, mean, and variance, extracted from the spatial domain to perform further tasks such as image segmentation or object detection.

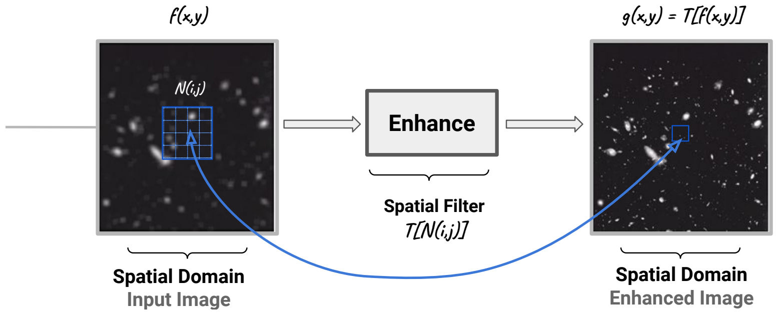

Fig. 3.3 Digital Image Processing in Spatial Domain. Example of image enhancement via spatial processing.

Convolution Operation. It involves sliding a matrix (kernel) over each pixel and its local neighbors in the image. Thus the kernel's size and values determine the transformation effect of the convolution operation, allowing various applications such as sharpening, edge detection, and DL tasks such as classification or object detection. With the proper padding and trimming operations, it produces a new image of the same size as the input, where each pixel is a weighted sum of its neighboring pixels. Mathematically, convolution is described as follows:

$$g(x,y) = w(x,y) * f(x,y) = \sum_{\delta x=-a}^{a} \sum_{\delta y=-b}^{b} w(\delta x, \delta y) f(x-\delta x, y-\delta y) \tag{3.2}$$

where $g(x,y)$ is the output image at position (x,y), $f(x,y)$ is the input image, $w(x,y)$ is the filter kernel at (x,y) positions. Every element of the filter kernel is considered by $-a \leq \delta x \leq a$ and $-b \leq \delta y \leq b$. The operator $*$ denotes convolution.

3.1.4 Digital Image in Frequency Domain

Digital images can be represented in the frequency domain through the Fourier transform. Within this domain, spectral properties such as sharpness, contrast and texture can be analyzed, hard to observe in spatial domain. Fourier transform depicts an image as a sum of complex sinusoids with different frequencies and orientations. In the frequency domain, high-frequency components correspond to edges and details in the spatial domain, while low-frequency components represent smooth regions and gradual intensity changes.

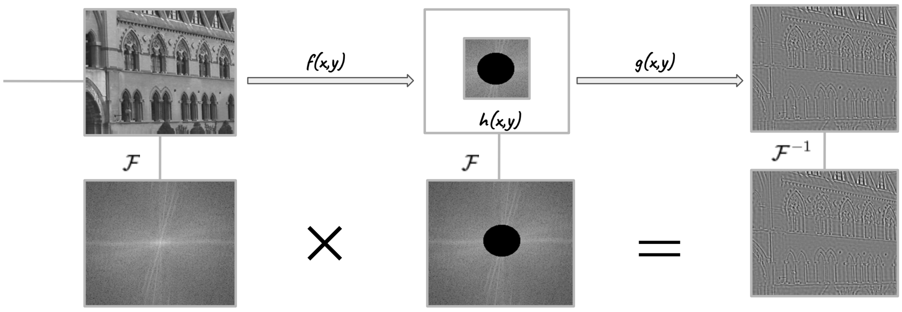

Fig. 3.4 Frecuency domain filtering operation: fourier transform, filter function and inverse fourier trnasform.

Discrete Fourier Transform. Digital images are discrete, so their Fourier Transform is also discrete. The 2D Discrete Fourier Transform (DFT) of an image $f(x,y)$ of size NxM is defined as:

$$F(u,v) = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x,y) e^{-i2\pi \frac{ux}{N} + \frac{vy}{M}} \tag{3.3}$$

where $F(u,v)$ is the frequency component at spatial coordinates (u,v) in the frequency domain. $f(x,y)$ is the pixel value at spatial coordinates (x,y) in the spatial domain.
Inverse Fourier Transform. The inverse Fourier Transform (IFT), illustrated in Fig.3.4, it takes a signal from the frequency domain back to the original spatial domain. For a two-dimensional grayscale image, the equation for the IFT is given by:

$$f(x,y) = \frac{1}{NM} \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} F(u,v) e^{i2\pi \frac{ux}{N} + \frac{vy}{M}} \tag{3.4}$$

where $f(x,y)$ is the reconstructed pixel value at spatial coordinates (x,y) in the spatial domain. N is the width of the image and M is the height.

3.1.5 Neural Networks and Deep Learning

Neural networks (NNs) consist of interconnected nodes, forming an acyclic graph. These connections have weights (θ) and a bias, with an activation function transforming input into decisions for hidden layers. Through multiple weighted hidden layers, data reaches the output layer for a solution. If the solution is not satisfactory per loss function, errors trigger θ updates using activation function gradients [64]. Furthermore, Convolutional Neural Networks (CNNs) are a specialized type of NNs for training on multidimensional data [65], which gained popularity since the AlexNet model outperformed ML-based models in 2012 ImageNet Large Scale Visual Recognition Challenge.

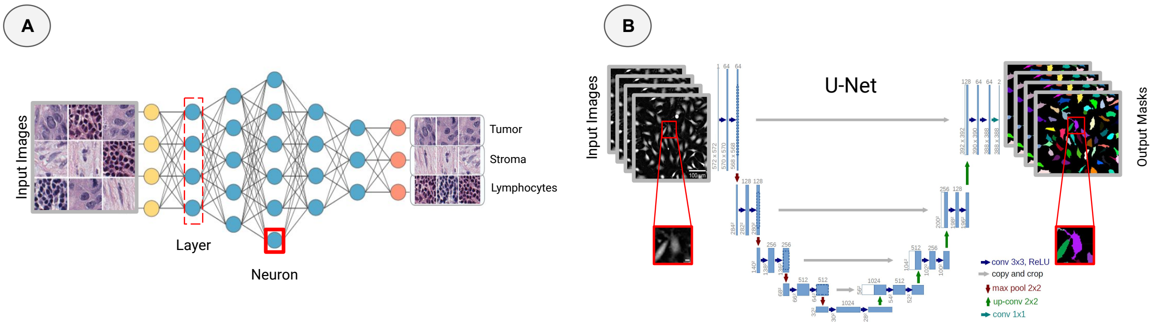

Fig. 3.5 Different neural network architectures. (A) Fully Connected Neural Network (FNN) where every node of each consecutive layer is connected. (B) Convolutional Neural Network (CNN) which learns kernels that capture the key features to represent an image.

Deep Learning (DL) has revolutionized biology by automating tasks and integrating complex data for reliable predictions [64]. Initially developed for computer vision, DL is currently applied to bioimage analysis[66] tasks such as cell segmentation, detection or classification (described in detail in Section.3.2). Its implementation benefits from high variability of images from different phenotypes, imaging modalities and acquisition settings. The use of Artificial Neural Networks (ANNs) in bioimage analysis dates to late 1980*s*, with the popularization of back-propagation algorithm [67]. However, it was massively adopted decades ago [68–70] for biomedical imaging, not for bioimage analysis until recent years [71–73]. Additionaly, ML enables computers to learn from data without explicit programming [74]. It includes supervised and unsupervised learning strategies. Supervised learning is the task of learning a function that maps an input to an output based on sample input-output pairs. It uses human-provided "ground truth" labels for model training, minimizing a loss function evaluated on a testing set [75, 5]. Unsupervised learning, including clustering and dimensionality reduction, and recently used in single-cell omics analysis [76], utilizes unlabelled input data to uncover patterns without human-provided examples.

The U-Net Revolution. U-Net [77, 78] is encoder-decoder initially developed for biomedical image analysis (detailed in Fig.3.5 (B)), which has encoding levels in the contracting path (encoder), a bottleneck and decoding levels in the expanding path (decoder). It was presented in 2015 at the International Symposium on Biomedical Imaging (ISBI). Since then, it has been extensively used for 2D and 3D cell segmentation becoming a powerful tool for bioimage analysis, whose effectiveness depends on the quality and quantity of training data, and tuning of network architecture.

3.1.6 Key Image File Formats in Light Microscopy

Microscopy image file formats are specialized file formats used to store and exchange data from various types of microscopes and manufacturers. The choice of proper microscopy file format often depends on the microscopy system, the software employed for acquisition, and the specific requirements of the analysis. It is important to consider the compatibility and metadata capabilities of each format to effectively manage and analyze their microscopy data. Fortunately, the open-source Bio-Formats [79, 80] library enables different formats to be read by many software such as Fiji or QuPath, and can be installed as a plugin for ImageJ. Even though it is written in Java, Bio-Formats can also be used within some Python applications. It is capable of parsing pixels and metadata for various formats, and writing to several formats.

File Format	Description
`.CZI` (Carl Zeiss Image)	Multidimensional,time lapse,Z-stacks,Multiposition experiments
`.ZVI` (Zeiss Vision Image)	HR image, 3x16-bit color and 16-bit,metadata and settings.
`.ND2` (Nikon NIS-Elements Data)	Metadata, annotations, time series, channels
`.LIF` (Leica Image File)	Multi-channels, metadata, time series, z-stacks
`.SCN` (Leica SCAN)	Pyramidal tiled BigTIFF with non-standard metadata
`.OIF` (Olympus Image Format)	Multi-file format including `.tif`,`.bmp`,`.txt`,`.pty`,`.roi`, `.lut`.
`.OIB` (Olympus Image Binary)	Compound file, storing OIF and associated files within one file
`.OME.TIFF` (Open Microscopy Environment)	Strengths of `OME-XML` (metadata) and `TIFF` (pixels).
`.OME.ZARR` (Open Microscopy Environment)	File format for cloud reading and writing image

Table 3.1 List of chief Microscopy-Related File Formats used in bioimaging.

3.2 Common Phases of a BioImage Processing and Analysis Pipeline

Advances in microscopy and imaging have increased complexity and volume of biological data, thus there is a growing need for sophisticated image analysis to automatically process these large and complex data. In current bioimage analysis pipelines, classical and DL algorithms synergize to extract meaningful quantitative information from biological images [81]. Another core aspect are the user-customizable tools, which unveil intricate processes. Also automation, since the amount of data generated by modern microscopes is staggering, being manual analysis often impossible. Therefore, reaching automation from the simplest batch processing to more complex routines, allows for handling large datasets efficiently.

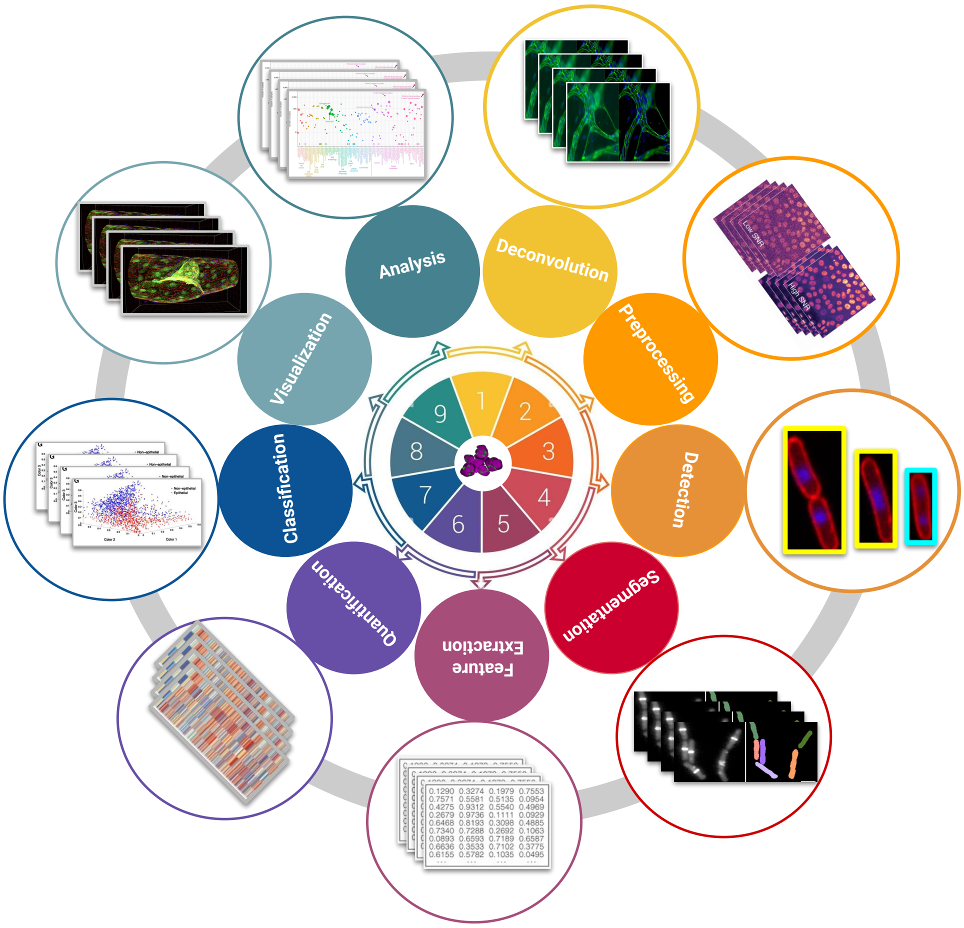

Fig. 3.6 Common Phases of a Bioimage Processing and Analysis Pipeline. Depending on the biological issue or the required application, not all steps may be always followed

DL techniques revolutionize bioimage analysis for tasks such as segmentation, detection, and classification [74]. Integrated correlated multi-modal imaging (CMI) provides a comprehensive view of biological structures [82] from different microscopy modalities. Hence dealing with data complexity demands efficient compression algorithms and standardization efforts to address format and metadata heterogeneity (Bio-Formats and OME [79, 80]). Yet, these challenges continue to reshape bioimage processing, with potential to revolutionize biology.

3.2.1 Image Restoration: Deconvolution

In optical microscopy, deconvolution (schematically shown in Fig.3.7) is a computerized inversion method to restore the original image from a blurred one. Deconvolution reduces out-of-focus blurring and the effects of random noise [22] during the image formation, compensating for microscope limitations. Deconvolution is highly effective in restoring 3D fluorescence microscopy from various imaging modalities [40, 83]. There are a variety of algorithms to perform deconvolution including: linear, iterative and blind methods. Linear deconvolution is suitable when the blurring process is well-defined and the PSF is known. Blind deconvolution is employed when neither the PSF nor the original image is known and aims to estimate both simultaneously. Iterative deconvolution assumes a known PSF and iteratively refines the image estimate using an optimization process to converge towards a desired solution. The choice of method depends on the specific characteristics of the image and the level of information available about the blurring process.

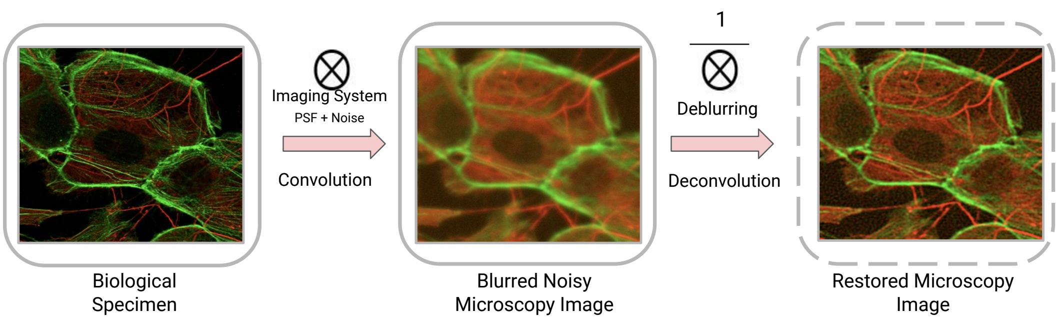

Fig. 3.7 Deconvolution operates by simulating the distortions that occur during imaging, and eliminating those distortions to approximate the appearance of the original sample.

Current DL methods [84–86] are promising in restoring details and resolving biological structures. CNNs create super-resolved images by training with many such pair of low-high resolution images [87]. To invert the convolution in the spectral domain, we require the spectrum of the denoised and degraded images, the noise spectrum ε, and the DFT of the PSF, known as the modulation transfer function (MTF).

$$\hat{F}(u,v) = F(u,v)H(u,v) + \varepsilon(u,v) \tag{3.5}$$

Neglecting the noise (ε), the most straightforward deconvolution method for recovering an initial (perfect) image from the degraded involves inverse filtering:

$$\frac{\hat{F}(u,v)}{H(u,v)} = F(u,v) + \frac{E(u,v)}{H(u,v)} \rightarrow F(u,v) \approx \frac{\hat{F}(u,v)}{H(u,v)} \tag{3.6}$$

using $1/H(u,v)$ as an inverse filter to remove degradation encounters several problems: indeterminate or infinite ratios due to zeros in the MTF and noisy data. Another theoretically solution is the Wiener filtering [40, 22], which minimizes the expected squared error between the restored and perfect images.

3.2.2 Preprocessing

Preprocessing is a crucial step in bioimage analysis, aiming to enhance image quality for reliable feature extraction and further analysis. It involves actions such as noise removal, geometric distortion correction and contrast enhancement, while preserving image details. Despite deconvolution can handle blurring caused by unstable imaging conditions, heterogeneous samples and technical limitations of microscopy system, there are external factors which can lead to image degradation, potentially biasing biological conclusions. Preprocessing normalizes and standardizes image features, ensuring data consistency and comparability. Current studies [88–90] strongly empathise the role of preprocessing when applying DL algorithms in microscopy images, as they remove the effect of noise, thus improving the efficiency of a model to generate accurate predictions. Accordingly, DL-based preprocesing outperforms classical methods due to their ability to handle noise variability with higher accuracy [91, 92]. However, choosing the appropriate the denoising method for a certain application relies on different factors such as the type and level of noise, the available computing resources, and the desired level of accuracy.

Image Denoising

Noise in fluorescence microscopy originates from limited resolution during acquisition, uneven background, out-of-focus light or properties of the fluorescent samples. The main noise sources are photon shot noise and detector noise. The measured signal x_i in Analog-to-

Digital Counts (ADC) can be represented as:

$$x_i = a\varphi(s_i) + \varepsilon_i \tag{3.7}$$

where $\varphi(s_i)$ denotes the shot noise-affected signal, a is the photon-to-ADC conversion factor, and ε_i represents detector noise.

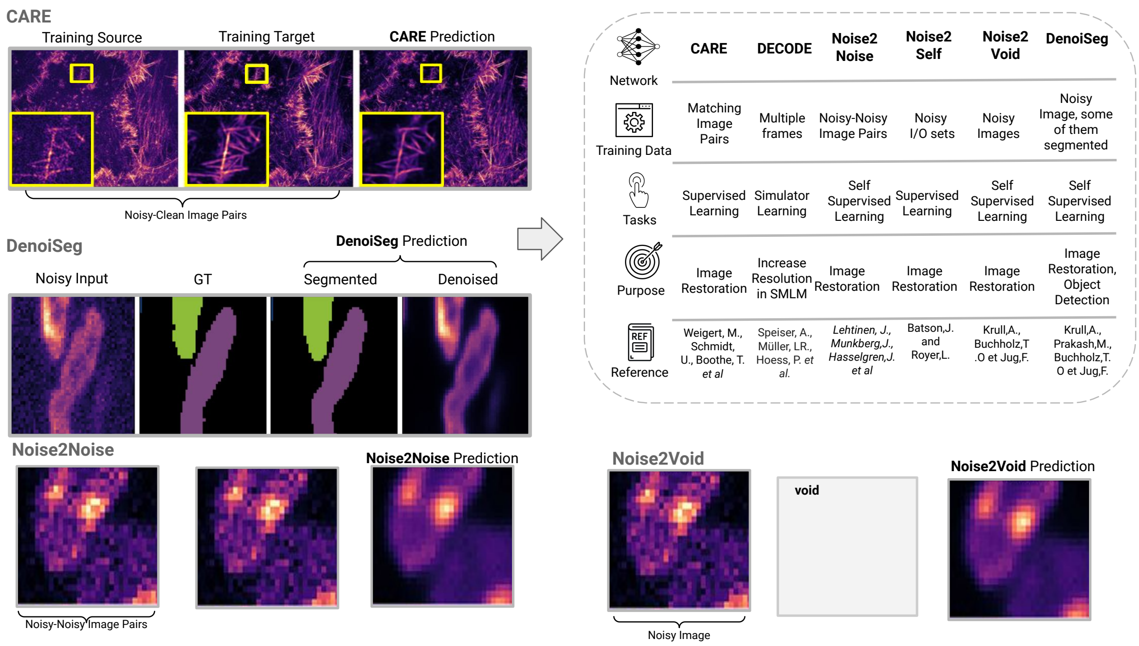

Fig. 3.8 Denoising and Image Restoration networks. Example of data generated using CARE, DenoiSeg, Noise2Noise and Noise2Void networks. Input and associated ground truth images used to train the network along with corresponding predictions are displayed.

Photon shot noise is related with randomness in photon emission and detection, being higher when fluorescence signal is weak (less photons emitted). Fluorophores emit photons stochastically, and the number of photons detected at a given time point follows a Poisson distribution centered around the underlying signal s_i. Likewise, detector noise (ε_i) exhibits a Gaussian distribution, it is inherent to electronics of imaging sensor and readout used [93]. Denoising methods rely on mathematical and signal processing techniques to reduce noise and restore the original signal. These approaches encounter challenges in preserving smoothness, edge protection, texture maintenance, and artifact avoidance [94]. Classical methods can be classified into linear filtering in which the output pixel value is a weighted sum of the neighboring pixel values (gaussian smoothing, mean, sobel-operator and laplace filter), non-linear filtering applying operations which can vary in a non-linear manner based on the pixel values in the vicinity (min, max, median, std...). Frequency domain methods involve manipulating frequency components as noise mainly resides in the high-frequencies, whereas image information is concentrated in the low-frequencies In this regard, the Wiener filter uses previously known spectral properties of image and noise to estimate and reduce it, and the butterworth filters are designed to pass or attenuate specific frequency components. Moreover, the wavelet-based methods decomposes the image into different scales and denoising is performed by thresholding and shrinking wavelet coefficients in these sub-bands [95]. Variational denoising formulates denoising as an optimization problem, incorporating a data fidelity term measuring noise difference and a regularization term enforcing smoothness [96, 97]. Finally, morphological methods (e.g., erosion, dilation...) can be used to reduce noise in binary images, and the top-hat and bottom-hat transform which highlight and extract small, bright details or dark features in the image, respectively. Although classical methods relying on theoretical knowledge of imaging systems to reach higher SNR[19], DL approaches have shown higher performance directly learning from complex relationships among noisy images and their corresponding ground truth [98]. Thus DL can be conceived as a sophisticated mathematical function which maps a noisy image to its clean version [93]. However, a challenge in supervised DL is the requirement of training ground truth images with minimal noise. DL methods such as CARE [99], DECODE [100], Noise2Noise [101], Noise2Self [102], Noise2Void [92] and DenoiSeg [92] are described in detail in Fig.3.8.

Image Registration

This is the procedure of finding a spatial deformation to spatially match two images (2D or 3D) [103] from same sample acquired under different conditions, imaging modalities or over time. Its goal is to find a function $g(x) : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ which maps coordinates from the target image I_t onto the source image I_s, so that $I_s(g(x))$ (a warped version of the source image) resembles $I_t(x)$ as much as possible. There are two transformation models: rigid/affine and non-rigid/elastic. The simplest one is the rigid model [104] characterized by translation (x, y, z) and rotation $(\theta_1, \theta_2, \theta_3)$ parameters and isotropic scaling (shown in Fig.3.9(B-D)). Rigid model preserver distances within the image and parallel lines. On the contrary, when more distortion is required such as shear (shown in Fig.3.9(E-F)), the transformation model is affine (shown in Fig.3.9(G)) having three scaling and three shearing parameters, it preserves parallel lines but not distances. Both rigid and affine models globally align pre-identified landmark features. While they are relatively robust against local minima, they are accuracy-limited due to local geometric difference is ignored [105].

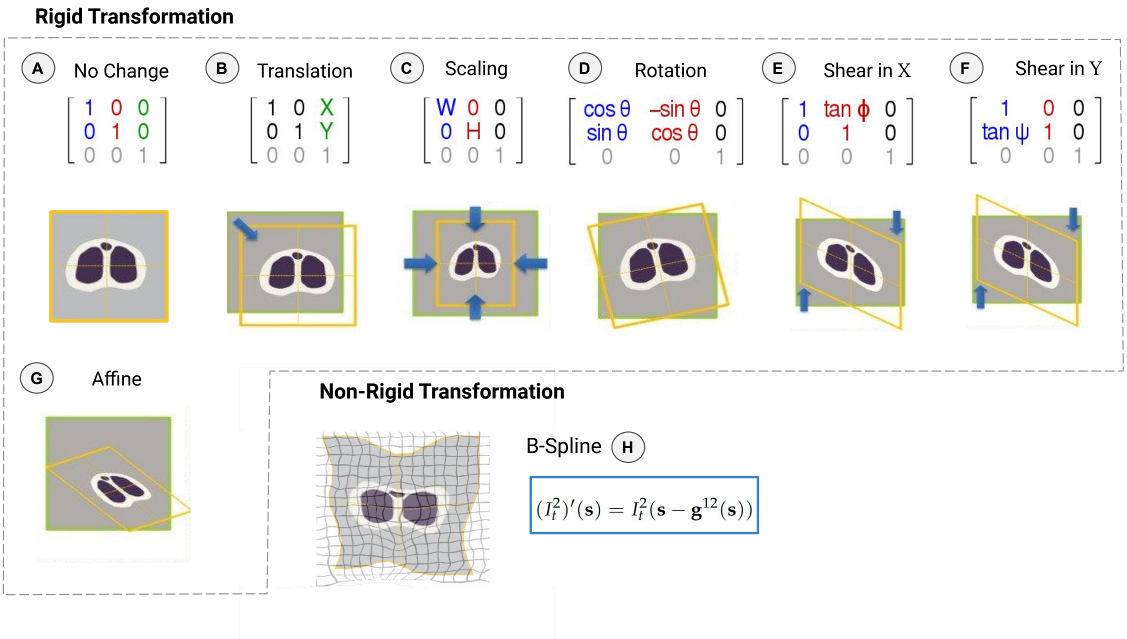

Fig. 3.9 Effect of Rigid and Non-Rigid Transformations. 2D Affine matrices contain 9 values, 6 relative to linear transformations of the X and Y coordinates. B-spline transformations require a set of control points and knot points.

The mathematical expression of affine transformation can be expressed as:

$$T_{affine} = \begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \\ \theta_{31} & \theta_{32} & \theta_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} \theta_{14} \\ \theta_{24} \\ \theta_{34} \end{pmatrix} \tag{3.8}$$

Conversely, when the deformation goes in different directions and magnitudes across the image, the transformation model is named elastic/non-rigid. This transformation model can deal with more sophisticated deformation. These elastic approaches includes elastic/non-rigid deformations such as thin-plate splines [106], which use a set of control points to estimate the transformation and B-spline transformations [107, 108], which use a set of control points and knot points to define a smooth deformation field.

In image registration, an objective function is defined to quantitatively assess the similarity of two aligned images. This criterion can be: landmark-based, firstly requiring the identification of corresponding homologous features as landmarks (points, lines, contours...) to be then mapped to each other, giving rise to the transformation model of two images; and intensity-based, which elastically align two images depending on intensity patterns [103]. For such, to optimize the objective function in a global o local manner, an optimization procedure is applied [105]. Hence the image registration is solved by iteratively searching for the parameters θ of a transformation model T which transforms a source (moving) image I_s into the reference space of target (fixed) image I_t. The best alignment is decided based on a distance measure d among the fixed and moving image. Registration can then be defined as $arg_{\theta} mind(T(I_s, \theta), I_t)$.

Bidirectional Image Registration based on Elastic Deformations represented by B-Splines Elastic deformation elastically simulate local deformations which can capture non-linear distortions and warping in the image. This approach warps local geometric features of a I_s (moving) for alignment with a I_t (fixed) image. Unlike rigid transformations, it can capture non-linear distortions, and it could be based on either a dense non-parametric model or a parameterized function model. In this regard, B-splines are piece-wise polynomial functions typically used to model both global and local deformation[107]. Since the B-spline is controlled locally, it is computationally efficient with many control points due to the following mathematical property: modifying a control point only affects its local neighborhood[109]. Moreover, b-splines are extremely useful to model the deformation field as they can be considered as a set of several functions (one per coordinate) which in turn are modeled by linear sum of weighted and shifted B-splines. The set of weights, which are called the B-spline coefficients, fully characterize the transformation. A deformation model based on B-splines is very versatile and can generate a large variety of nonlinear elastic deformations, while remaining easy to handle[107]. Thus elastic and consistent image registration based on B-splines becomes more and more popular since its superiority in the transparency, applicability, as it high smoothness and continuous transformation with high topology preservation[109]. The "direct" transformation (from I_s to I_t) is performed, in which I_s is elastically deformed to look as similar as possible to I_t, while simultaneously, the "inverse" transformation (from I_t to I_s) is also being computed. Therefore, a pseudo-invertibility of the final deformation is provided. By reducing the likelihood of being trapped in a local minimum, this approach enhances the registration process and allows for simultaneous registration of any number of images. The idea of elastic registration using vector-spline regularization [107] is known as consistent registration [110]. With this algorithm the energy functional presented in [107] is extended into a new functional which incorporates a factor of the deformation field consistency. Furthermore, it simplifies the search for the optimum deformation and allows registering with no information about landmarks or deformation regularization[111].

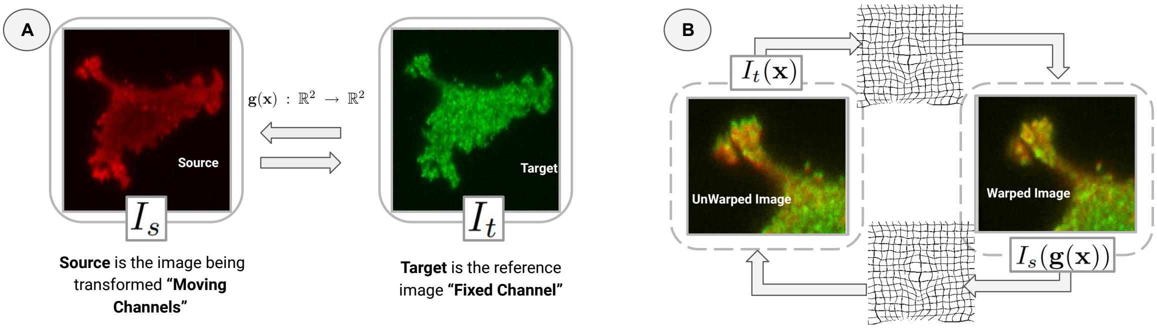

Fig. 3.10 (A) Bidirectional image registration based on elastic deformations represented by B-splines. Deformation Estimation from Image Registration based on Elastic Deformations represented by B-Splines. (B) A warped version of the source image ($I_s(g(x))$) resembles $I_t(x)$ as much as possible.

The algorithm relies on minimizing an energy functional, consisting of components: the dissimilarity between the source and target images in both directions (E_{img}), the optional landmark constraint (E_μ), a regularization term encompassing both divergence and rotation ($E_{div} + E_{rot}$), and an energy term (E_{cons}) representing the geometric consistency between bidirectional elastic deformation (from I_s to I_t and from I_t to I_s). As a result, the energy function now comprises four terms, as follows:

$$E = w_i E_{img} + w_\mu E_\mu + (w_d E_{div} + w_r E_{rot}) + w_c E_{cons} \tag{3.9}$$

where w_c is the specific weight given to the new consistency term.

Similarly, the deformation field is defined as a linear combination of B-splines by following:

$$\mathbf{g(x)} = g(x,y) = (g_1(x,y), g_2(x,y)) = \sum_{k,l \in \mathbb{Z}^2} \begin{pmatrix} c_{1,k,l} \\ c_{2,k,l} \end{pmatrix} \beta^3(\frac{x}{s_x} - k)\beta^3(\frac{y}{s_y} - l) \tag{3.10}$$

where s_x and s_y are scalars (sampling steps) controlling the degree of detail of the representation of the deformation field.

The algorithm implemented in Paper *III* (detailed in *Apendix C*) provides invertible deformation field as it extends unidirectional registration to bidirectional by performing a simultaneous registration of two images in a single computation (as shown in Fig.3.10 (A-B)).

3.2.3 Object Detection and Image Segmentation

A common strategy for object detection relies on considering objects as clusters of bright pixels. One method is to search the local maxima, pixels with higher intensity than neighbors, to identify objects as peaks. Noise leads to false detections and challenges detection strategy, but tools such as ImageJ detect local maxima by considering a pixel as a local maximum if none of its eight neighboring pixels have a higher intensity [112]. The Laplacian of Gaussian (LoG) filter (detailed described in Fig.3.11), yields precise and robust results. It is sensitive to bright and roundish objects of a specific size, less affected by noise. To enhance detection, the LoG-filtered image intensity can be used as a quality metric in which thresholding removes undesired peaks. Software packages such as FeatureJ, SpotTracker, and TrackMate offer LoG-based peak detection [113–115]. Furthermore, fitting approaches are key for object detection when objects lack clear/crisp features or match fitting functions [116]. Image processing simplifies complex raw images into intermediate images with reduced content, suitable for fitting. A cell detection method clusters pixels into supervoxels before Gaussian mixture fitting [117].

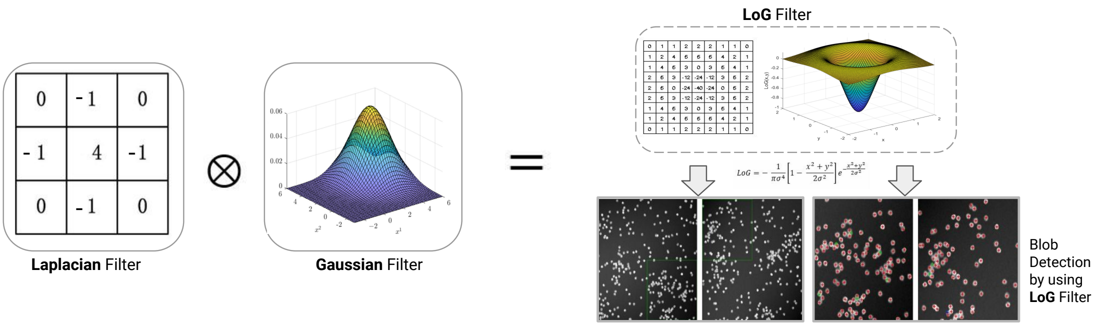

Fig. 3.11 Laplacian of Gaussian (LoG) Filter. Input Image needs to be smoothed (by convolution with the Gaussian filter) then, the smoothed image needs to be convolved with the 3x3 Laplacian filter to obtain the output image.

Image segmentation is crucial as it splits images into foreground and background. It enables further bioimage analysis tasks such as object counting, distribution, shape, recognition, tracking, or region removal [118]. Super-resolution fluorescence microscopy and computer vision have enhanced accuracy and efficiency of cell segmentation, since it plays a crucial role, enabling analysis of cell count, type, division, and shape [119]. Segmentation methods include semantic segmentation (Fig.3.12(B)), which classifies each pixel into a specific category or class, allowing for the identification of objects and regions based on their semantic meaning. Also, instance segmentation (Fig.3.12(C)), which not only categorizes pixels into object classes but also distinguishing individual instances of objects within the same class, enabling unique identification of each object instance. In this regard, Otsu's method [120] is a typical approach for segmentation in biology, which uses gray threshold to separate foreground and background pixels by minimizing intra-class variance [121].

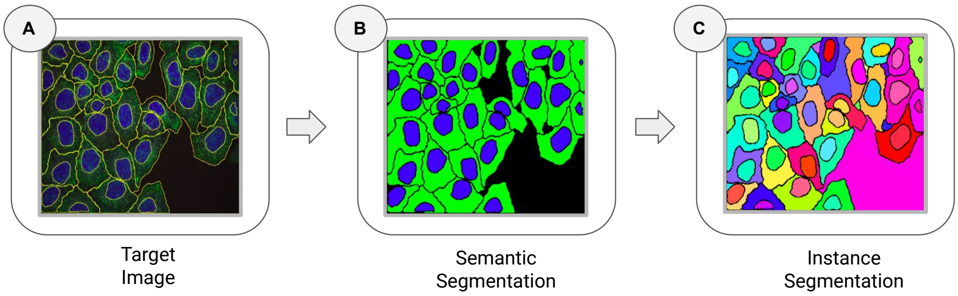

Fig. 3.12 Graphical representation of differences among (B) semantic segmentation and (C) instance segmentation.

Classic segmentation algorithms emphasized gray-level similarity within regions and discontinuity among regions. Color image segmentation identifies similar pixels and merges regions [122]. Software such as ImageJ, CellProfiler, CellCognition automate [123]these steps. Moreover, automated thresholding algorithms reduce shape errors by using image content to determine thresholds. Morphological operations (dilation, erosion, closing, opening) clean nosiy masks and smooth contours while maintaining size [124–126]. Watershed approach divides images into catchment basins based on markers but over-segments non-round/elongated objects [127]. Complex algorithms such as deformable contours (snakes or level sets) iteratively adjust an initial contour to outline object boundaries using partial-deritvative equations and shape constraints. Fiji and Icy implement plugins for deformable contours: E-Snake [128] for Fiji, and various plugins for Icy [129–131]. ML-based segmentation tools, such as Trainable Weka Segmentation utilizes the Weka toolbox [132, 133]. Icy's Rapid Learning plugin employs RapidMiner, and Texture Segmentation combines color and texture features. ML methods such as clustering groups similar pixels in images, Pixels in an $M \times N$ image are represented as vectors $P = (x, y, I(x, y))$, with (x, y) as pixel locations and $I(x, y)$ as feature vectors. On the other hand, template matching assigns a class to each pixel by finding the most similar template based on pixel values. The distance $||I(x, y) - r_{c,k}||$ is evaluated, where $I(x, y)$ is the pixel value and $r_{c,k}$ is a template for class c. The recognition result is the class c with the minimum distance [134].

DL-based methods have revolutionized object detection and segmentation. U-Net, adapted for instance segmentation, predicts cell interiors, edges, and background effectively. Mask R-CNN [135] and and you-only-look-once (YOLO) [136] adaptations succeeded in nuclei segmentation and detection, respectively [137]. Building upon U-Net, StarDist [138] improves nuclei segmentation by predicting star-convex contours, aiding overlapping nuclei separation in 2D/3D images. SplineDist [139] extends this to segment more complex shapes. Benchmark datasets, such as the 2018 Kaggle Data Science Bowl dataset [140], advances 2D nuclei segmentation. However, cell membrane segmentation is tougher due to varied cell morphology, lacking benchmarking datasets. To overcome these challenges, Cellpose [141] uses U-net models trained on vast microscopy datasets, predicting spatial gradients for 2D/3D data [142, 143].

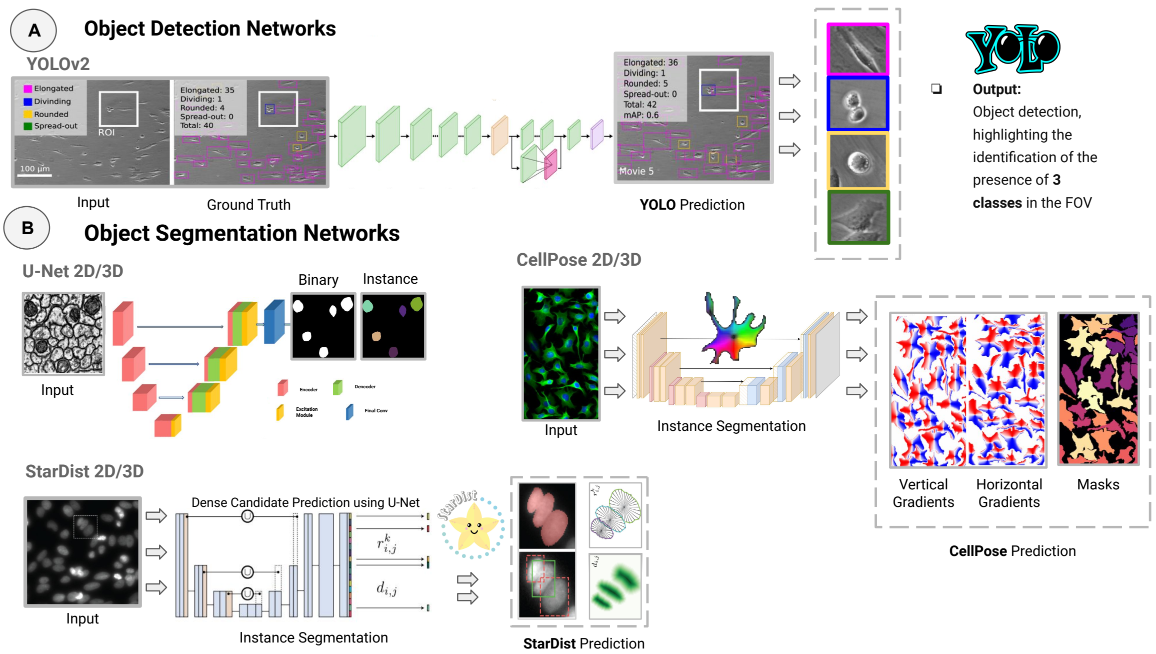

Fig. 3.13 Object Detection and Segmentaion Networks. (A) Examples of using ZeroCostDL4Mic YOLOv2 notebook to detect and identify cell shape classification from cell migration bright-field time-lapse dataset. (B) Examples of using U-Net, CellPose and Stardist networks for instance cell segmentation.

3.2.4 Feature Extraction

Since the early 1960*s*, advanced computing enables automated feature extraction [144]. Before image classification, relevant features are extracted from biological images [145], crucial for describing pixels, voxels, and higher-level objects. Feature extraction captures meaningful information representing specific image patterns, applied to cells, sub-cellular structures, or tissue regions [146]. Examples (details in Table.3.2) encompass cellular morphology, organelle structures, and intracellular biomolecule levels [146]. Intensity-based features quantify intensity regularity within regions, offering insights into distribution and patterns. These metrics are computed from histograms [147].

Feature	Type	Mathematical Definition
Mean;Variance	Intensity	$\mu = \frac{\sum I(i,j)}{N}; \quad \sigma^2 = \frac{\sum (I(i,j)-\mu)^2}{N}$
Skewness;Kurtosis	Intensity	$SK = \frac{\sum (x-\overline{X})^3}{(n-1)SD^3} \quad KT = \frac{\sum (x-\overline{X})^4}{(n-1)SD^4}$
Haralick Features	Intensity	
Gabor Filters	Intensity	$g(x,y;\lambda,\theta,\psi,\sigma,\gamma) = exp(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}).cos(\frac{2\pi x'}{\lambda}+\psi)$
Area;Perimeter	Shape	$A_O = \sum_{(x,y)\exists O} B(x,y); \quad P = A - A_{eroded}$
Circularity;Eccentricity	Shape	$C = \frac{4\pi.A}{P^2} =; \quad E = \sqrt{1-\frac{L_{Minor}}{L_{Major}}}$
Aspect Ratio;Solidity	Shape	$AR = \frac{L_{Major}}{L_{Minor}}; \quad S = \frac{A_O}{A_{convexhull}}$
Ferret Diameter	Shape	$FD = max_\theta[max_{p\exists P}(p.cos(\theta)) - min_{p\exists P(p.cos(\theta))}]$
LBP GLCM	Texture	
Wavelet-based Features	Texture	$WF = \sum_{n=1}^{N}\sum_{i=1}^{W}\sum_{j=1}^{H} \lvert DWT(i,j,n)\rvert^p$
SGLD	Texture	7

Table 3.2 Intensity, shape and texture based features together with corresponding math equations.

In addition, texture features assess intensity regularity, capturing properties such as smoothness or roughness [148]. Gray Level Co-occurrence Matrices (GLCM) define texture via statistical pixel pair relationships at specific distances and angles [149]. Local binary patterns (LBP) extract local contrast and texture features such as uniformity, contrast, and entropy [148]. Densitometric features, from Spatial Gray Level Dependence matrix (SGLD), include entropy, energy, and correlation [150, 151]. Shape features quantify area and geometric features (Ferret diameter, eccentricity...) [152]. These features are vital in cell classification, cancer diagnosis, or tissue characterization, revealing abnormalities, and understanding biological processes. Skeletonization reduces an object to its thinnest representation, preserving topology, and extracts branch points, end points, and branch lengths to evaluate connectivity [111]. Spatial features include distance metrics (nearest and farthest neighbors) and spatial moments indicating pixel intensity distribution relative to position, size, and orientation. Examples are the center of mass, a weighted average of pixel intensities and coordinates, and centroid, an arithmetic mean of pixel coordinates, provide more robust object location [153].

3.2.5 Feature Selection and Classification

Image classification is a fundamental task in computer vision consequently, in bioimage analysis, assigns images to predefined classes using distinctive features. It involves dividing data into training to define classification rules and testing subsets to evaluate performance. To enhance efficiency, feature selection identifies most informative features, by simultaneously eliminating irrelevant, reducing data dimensionality and improving classifier performance. In the past, assigning objects to specific pre-defined categories was tackled through the extraction of manually-engineered features, it was time-consuming and biased. Nowadays, these features, along with predefined class labels, are used to train ML classifiers as k-nearest neighbors (KNN), support vector machines (SVM), random forests (RF), and decision trees (DT). Non-parametric algorithms such as KNN, predicts based on similarity of new instances to labeled instances. SVMs maximize class separation in high-dimensional data. Also, RF combines decision multiple decision trees trained on random sets of data for improved generalization, while DTs partition feature space and learn decision rules for class prediction [154, 146]. DL excels in image classification, surpassing traditional methods in accuracy and efficiency. Some use DL for embryo quality assessment based on Google's Inception-V1 architecture [155], others for assessing microscopy focus quality regardless the miscroscopist [156]. Lastly, DL showed higher performance in cell classification, sub-cellular pattern recognition [157], protein localization from yeast and humans [158, 159]. Yet, recent evaluations suggest DL does not always outperform classical methods [160], attributed to limited training data, which transfer learning can mitigate [98]. Furthermore, image classifier evaluation is crucial for performance assessment. Metrics such as accuracy, precision, recall, F1-score, and ROC curves measure effectiveness.

3.2.6 Common Analyses in BioImage Analysis

Cell-Type Analysis

In single-cell data analysis, a typical procedure involves the annotation of cells according to their phenotype., especially in high-throughput experiments [161, 146]. Recent advancements have facilitated the evaluation of treatment conditions, enabling the systematic assessment of cell morphologies. Therefore, assessing the impact of treatments involves measuring numerous morphological features and comparing changes between conditions [162]. Cell-Type analysis involves the automatic or semi-automatic identification and categorization of cells based on their morphological, structural, or functional characteristics observed in microscopic images. Cell-Type identification has diverse applications in various fields of life sciences, including cancer diagnosis, neuroscience, drug discover, stem cell research or immunology. [163, 144]. Defining cell-types was typically subjective and relied on manual annotations, where experts visually examined images and manually measured specific features. This approach was time-consuming, subjective, and limited in scalability. Prior to cell-type analysis, acquired images often undergo preprocessing steps to enhance the visibility of cell structures. Then cell segmentation is a critical step where single cells are delineated from the background and adjacent cells. Subsequently, relevant features are extracted. These features can include shape descriptors (area, perimeter, aspect ratio, roundness...), texture analysis, intensity statistics, spatial measurements (spatial distribution, relationships among cells and their neighbourhood...), or temporal features (cell migration, cell dynamics...). However, selecting appropriate features and handcrafting them for different cell types is still challenging. Decades ago, this process aimed to assign specific labels or annotations to cells, by utilizing specific biomarkers or proteins expressed by cells to identify their type or state. In recent years, DL approaches have shown significant promise in overcoming cellular heterogeneity, varying cell shapes and sizes, overlapping cells, and complex cell structures. Specifically, ML supervised algorithms, allow for training classifiers using the information contained in the pre-defined markers to detect cells of interest. Once cell features are extracted, classification algorithms are employed to categorize cells into different predefined classes. Unsupervised techniques, such as clustering are used to identify natural groupings cells based on their features (similarity in marker expression or by their proximity in low dimensional space) without prior class labels.

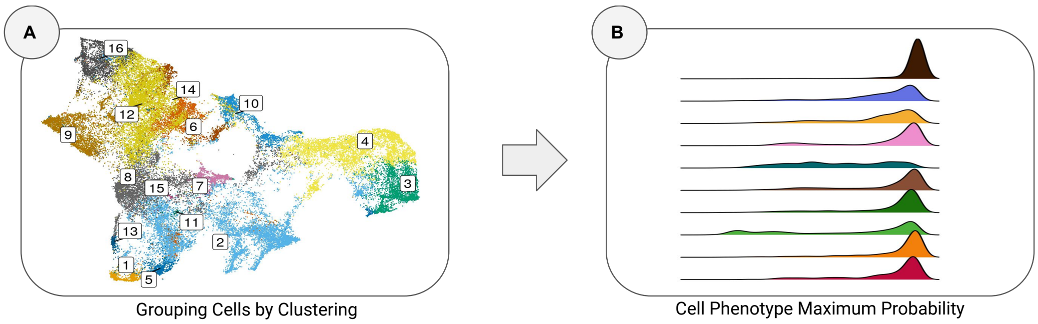

Fig. 3.14 Example of cell type analysis by using unsupervised clustering approach. (A) Expression thresholding of sixteen clusters or cell types. (B) Distribution of maximum probabilities, each cell is assigned to the class with highest probability.

> The approach presented in Paper *I* (detailed in *Appendix A*) provides a semi-automated solution to carry-out cell-type analysis and further user-customizable classification, easily implementable in most of light microscopy facilities' daily routines.

Colocalization Analysis

Fluorescence microscopy aids in studying spatial arrangements and protein interactions through colocalization analysis, revealing cellular functions and mechanisms [164]. However, it is not suitable for molecular interactions at super-resolution light microscopy ($\sim 70\breve{}90nm$). Colocalization is best for determining if two molecules/proteins associate with the same structures, subnuclear structures, or membrane domains, though fluorescence overlap does not guarantee it due to resolution limits. It is assessed visually, quantitatively, and statistically using co-occurrence (spatial overlap of probes) and correlation (distribution within/among structures), but careful method selection is crucial for proper analysis [165].

Methods to Quantify Colocalization Colocalization measures distribution of two molecules, capturing co-occurrence and correlation. Co-occurrence checks molecule presence together, while correlation measures similarity in concentration variation [166]. Pixel-wise colocalization compares pixel intensity among two channels, generating a scatterplot. Correlation degree is quantified by a coefficient. However, in super-resolution microscopy, pixel-wise matching struggles to demonstrate positive spatial correlation due to enhanced resolution causing limited overlap in closely correlated proteins. Therefore, image resolution and scale impact on results and interpretation.

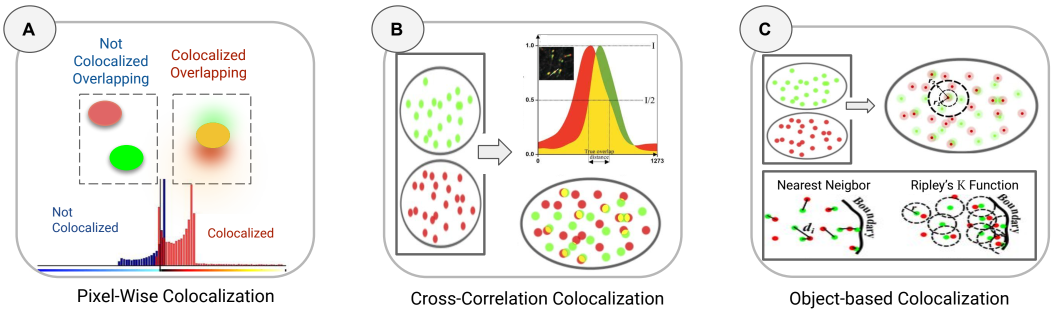

Fig. 3.15 Common methods to quantify colocalization. (A) Pixel-Wise Methods. (B) Cross-Correlation Methods and (C) Object-based Colocalization of Single-Molecule Localization Microscopy.

Pixel-Wise Colocalization Method	Mathematical Definition
Pearson Correlation Coefficient(PCC)	$PCC = \frac{\sum_i (R_i - \overline{R}).(G_i - \overline{G})}{\sqrt{\sum_i (R_i - \overline{R})^2 . \sum_i (G_i - \overline{G})^2}}$
Mander's Overlap Coefficient(MOC)	$MOC = \frac{\sum_i (R_i . G_i)}{\sqrt{\sum_i R_i^2 . \sum_i G_i^2}}$
Colocalization Coefficients m_1, m_2	$m_1 = \frac{\sum_i R_{i,colocal}}{\sum_i R_i} \quad , \quad m_2 = \frac{\sum_i G_{i,colocal}}{\sum_i G_i}$
Mander's Colocalization Coefficients (MCC) M_1, M_2	$M_1 = \frac{\sum_i R_{i,colocal}}{\sum_i R_i} \quad , \quad M_2 = \frac{\sum_i G_{i,colocal}}{\sum_i G_i}$
Overlap Coefficients k_1, k_2	$k_1 = \frac{\sum_i R_i . G_i}{\sum_i (R_i)^2} \quad , \quad k_2 = \frac{\sum_i R_i . G_i}{\sum_i (G_i)^2}$

Table 3.3 Mathematical definition of common pixel-wise methods to evaluate colocalization.

Pixel-Wise Colocalization Methods. Illustrated in Fig.3.15(A). Quantifying dual-color image correlation uses coefficients such as Pearson Correlation Coefficient (PCC) and Mander's Overlap Coefficient (MOC) (Table 3.3). PCC measures linear association, while MOC improves PCC by considering mean intensity variations [167, 168]. m_1 and m_2 show channel impact, both 1 means perfect colocalization. MCC notes pixel of interest contribution [169]. k_1 and k_2 split colocalization into two parameters, discerning each antigen contribution.

Colocalization by Cross Correlation Function. Intensity-based colocalization often uses Cross-correlation function (CCF) to assess channel correlation. CCF has two colocalization approaches: spatial and temporal. Spatial checks distance-based CCF for providing information about spatial overlap of molecules or proteins. It needs specific imaging, where one channel is shifted relative to the other, generating correlation curve as a function of distance. This method can detect spatial relationships even without signal overlap, great for super-resolution, but computationally intensive. Conversely, temporal uses time-based CCF, tracking fluorescence changes for molecular interaction info over time. This method needs time-lapse, not available for single-frame or fixed samples. This bunch of methods are illustrated in Fig.3.15(B).

Object-based Colocalization Method	**Mathematical Definition**
Coordinate-based Colocalization	$D_{x_i,x}(r) = \frac{N_{x_i,x}(r)}{N_{x_i,x}(R_{max})} \cdot \frac{R_{max}^2}{r^2} \quad and \quad D_{x_i,y}(r) = \frac{N_{x_i,y}(r)}{N_{x_i,y}(R_{max})} \cdot \frac{R_{max}^2}{r^2}$
Ripley's K-Function	$K_{ij}(r) = \frac{A}{N_i N_j} \sum_{i=1}^{N_i} \sum_{j=1}^{N_j} \frac{I(d_{ij}<r)}{w_{ij}}$

Table 3.4 Mathematical definition of common object-based methods to evaluate colocalization.

Object-based Colocalization of Single-Molecule Localization Microscopy (SML). Spatial statistics in SML microscopy quantify molecular associations without spatial overlap [170]. Malkusch et al. [171], introduced a coordinate-based colocalization, calculating two functions ($N_{x_i,x}(r)$ and $N_{x_i,y}(r)$, Table 3.4). The first counts x molecules (channel 1) within radius r of a given localizationx_i. The second counts y molecules (channel 2) within same radius r and maximum distance R_{max}. These functions assign correlation to each molecule in SML image, reflecting spatial association among two molecules. Ripley's K-function [172] is a key tool, measuring interaction distance between two molecules. In this context, A is imaging area, N_i and N_j are molecules localized in each channel, $I(d_{ij} < r)$ is 1 if $d_{ij} < r$, 0 otherwise. This bunch of methods are illustrated in Fig.3.15(C).

Insights into Tracking Analysis

In recent years, image recording and storage have enhanced experiments, enabling observation of cellular activities and aiding biological applications. Microscopes play a crucial role in visualizing these time-lapse images and tracking objects over time. Various methods and tools, employing advanced algorithms and deep learning, have been developed for single particle tracking (SPT), finding broad applications in stem cell viability, cell dynamics, and trajectory tracking.

Methods for Single Particle Tracking Jaqaman et al. [173] developed an algorithm for linking segmented particles over frames using a linear assignment problem (described in detail in Section 3.2.6). Yang et al. [174] proposed a probability-based framework with foreground and background markers for particle detection and a multiple mode filter for motion modeling. Meijering et al. [74] emphasized the challenges in detecting and tracking small particles in microscopy images and the need for global linking strategies. Various tools have been proposed for single particle tracking (SPT), including ClusterTrack, ManualTracking, MTrackJ, Mtrack2, and U-track. Vallotton et al. [175] introduced Tri-track, a software which simplifies SPT tasks using a graph structure. Chenouard et al. [176] proposed a Bayesian model and multiple hypothesis tracking algorithm for SPT in microscopy images. Shuang et al. [177] discussed the difficulties in quantitative analysis of SPT data and the need for faster and more reliable approaches, including GPU acceleration. Liang et al. [178] presented a SPT method for managing trajectories, solving data association problems, and handling pseudo-split/merged particles. Chenouard et al. [179] organized a competition to compare SPT algorithms and ML models, highlighting the benefits of multi-frame and multi-track optimization schemes. Jaiswal et al. [180] proposed an SPT approach based on multi-scale detection and two-step multi-frame association. Smal et al. [181] compared data association techniques for SPT, finding that multi-frame techniques generally outperform two-frame techniques. Furthermore, Tinevez et al. [115] developed TrackMate, an open-source tool for SPT with a user-friendly interface, allowing developers to create their own algorithms.

In the pre-DL era, the first international competition for tracking methods sparked the idea of DL-based methods [179]. Tracking still presents some challenges regarding cell stages which are being overcome by the successful incorporation of DL [182]. DL models have been trained to classify cell cycle stages and identify cell state trajectories from single-cell data [183]. While DL approaches show promise in cell division, classical ML methods trained with smaller datasets remain a competitive alternative in cell identification [184]. Efforts have been made to train DL models which utilize information from surrounding frames to identify cell matching pairs [98], and simulations have been used to build large training sets with less human intervention [185]. Tools such as DeepLabCut enable automated tracking of points on organisms with minimal manual annotations [186]. Despite progress in DL-based SPT, the challenge lies in creating end-to-end solutions [187]. Platforms like ZeroCostDL4Mic [188] provide accessible DL models for tracking, with compatibility to existing tools such as TrackMate [115] compatible with DL segmentation models.

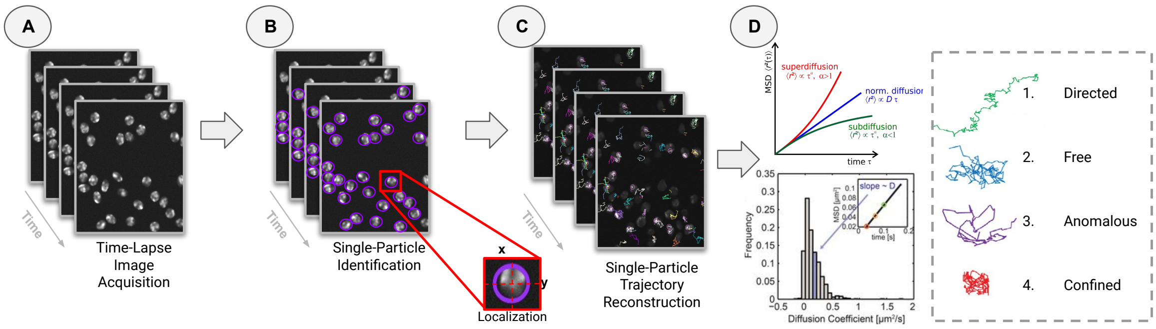

Fig. 3.16 Workflow to perform SPT and subsequent analysis of diffusion for motion classification. (A) After Acquisition of time-lapse data sets. (B) Localization, detection and identification of single particles over time. (C) Single particles are linked to build trajectories and features extracted. (D) Resulting trajectories are characterized and the type of motion evaluated by applying quantitative analysis of diffusion, MSD and MSS slope.

Solving Particle Tracking as Linear Assignment Problem The linear assignment problem (LAP) is a fundamental task in SPT analysis (showed in Fig. 3.17), aiming to establish correspondences among particles detected in consecutive frames. The LAP problem can be formulated as follows: Given a set of particles detected in frame t and another set of particles detected in frame $t+1$, the task is to find the most likely matching pairs of particles over frames which minimize a specific cost metric. The Jaqaman approach [173] is a widely used for solving the LAP problem. Hence it introduces the tracking analysis by tackling the primary challenges faced in SPT: high particle density, particle motion heterogeneity, temporary particle disappearance and particle merging/splitting.

The LAP algorithm addresses these challenges by linking particles among consecutive frames and linking resulting track segments to form complete trajectories. Both steps involve solving global combinatorial optimization problems, which determine the most probable set of particle trajectories throughout the entire sequence. SPT goes beyond particle detection and localization; it focuses on establishing correspondence between particle in a sequence of frames. In Multiple-Hypothesis Tracking (MHT), all possible particle paths within the expected behavior bounds are constructed based on the given particle positions in each frame.

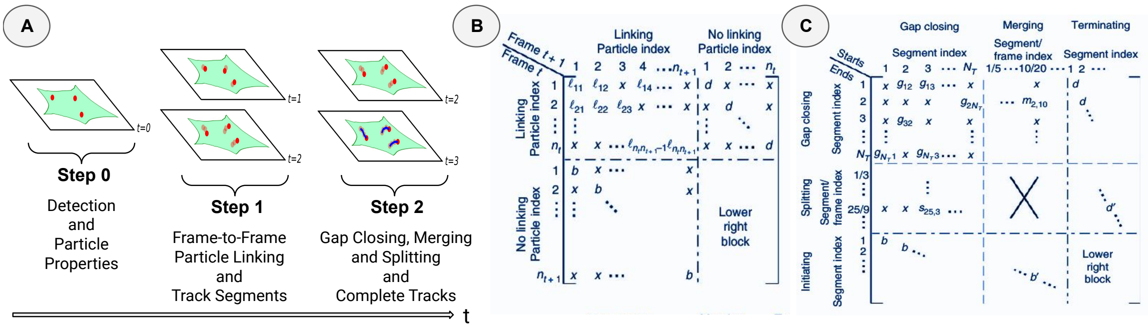

Fig. 3.17 Tracking particles through spatially and temporally global assignments involves the following steps. (A) Creating tracks from an image sequence involves detecting particles in each frame (step 0), linking particles across consecutive frames (step 1), and subsequently closing gaps while capturing merging and splitting events among the initial track segments (step 2). (B) A cost matrix is used to manage particle assignments between frames. (C) Another cost matrix is used to oversee the process of closing gaps, merging, and splitting.

The solution involves selecting the largest set of paths without conflicts, ensuring global optimally in both space and time. Nevertheless, this approach is computationally demanding. In the LAP framework, each potential assignment (partial assignment in the first step and track segment assignment in the second step) is assigned a cost C. The goal of solving the LAP in each step is to identify the combination of assignments with the minimum sum of costs. To handle cases with missing or false detections, the Jaqaman approach incorporates gap closing, merging, and splitting steps where six potential assignments were in competition:

Type of Event	**Description**
Gap closing g	Link the end of one track segment to the start of another
Merge m	Link the end of one track segment to a middle point of another
Split s	Link the start of one track segment to a middle point of another
Termination d	End of a track segment does not link to anything
Initiation b	Start of a track segment does not link to anything
NOT merge or split d' and b'	Middle points introduced for merging and splitting do not link to anything

Table 3.5 List of potential assignments in competition through linear assignment problem in the Jaqaman approach.

The LAP framework is independent to dimensionality and particle motion types. It is also not dependent on the physical nature of the particle (single molecule, molecular assembly, or organelle), except for the choice of a suitable particle-detection method. However, the cost function needs to be customized for each specific tracking application.

Trajectory Classification based on Motion Type

Understanding the motion behavior of particles within cells is essential for unraveling cell processes and protein cell entry mechanisms. Thus single particle needs to be imaged and their trajectories reconstructed to gain insights into cellular dynamics, migration patterns and interactions. Motion classification offers a powerful approach to categorize different types of particle motion, while provides information about their underlying biology in heterogeneous environments. The diffusion characteristics of trajectories can unravel distinctive motion patterns, and depending on it, particle movements can be categorized into four basic motion types [189]: free diffusion, anomalous diffusion, confined diffusion and directed motion. Free diffusion occurs when particle movements are totally unrestricted, whereas directed diffusion is an active process which can become evident when small corpuscles are transported by molecular machines along micro-tubules [190]. Confined diffusion is observable for trapped particles or particles whose free diffusion is confined by cytoskeletal elements [191]. On the other hand, anomalous diffusion is commonly traced back to the macromolecular crowding in the interior of cells, but its precise nature is still under discussion [192].

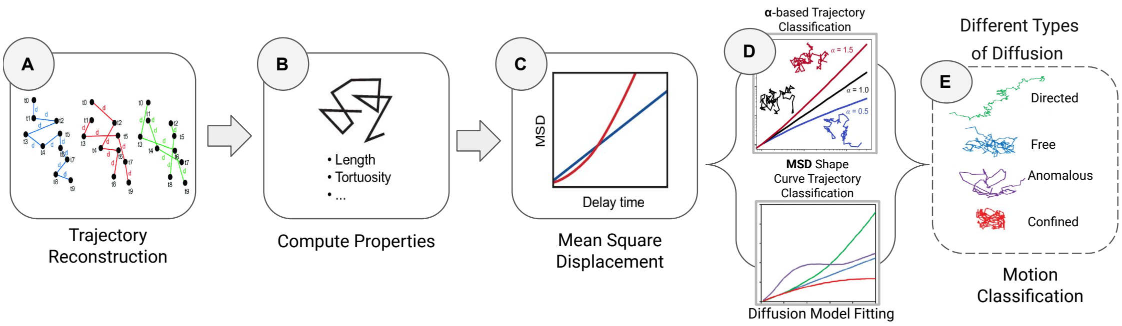

Fig. 3.18 Motion Type Classification of Trajectories. (A) After SPT, trajectories are reconstructed. (B) Trajectory features are computed. (C) MSD of each particle is computed. (D) MSD curve fitted through data follows one of these models, α is computed to disclose motion type for Brownian motion. (E) Motion Classification

A trajectory represents the number of N consecutive 2D positions of a particle $\mathbf{r}_j = (x_j, y_j)$ recorded with a constant time interval Δt over a period of time $T = (N-1)\Delta t$. A change in position from $\mathbf{x}_j$ to $\mathbf{x}_{j+1}$ is called a step, whose length is defined as the euclidian norm $|\mathbf{x}_j - \mathbf{x}_{j+1}|$. A subtrajectory is a part of a trajectory. The mean squared displacement (MSD) provides valuable insights into the characteristic diffusion behavior exhibited by the particles. The MSD of the particle j with time lag $n\Delta t$ is mathematically described as:

$$MSD_j(n\Delta t) = \frac{1}{N_j - n} \sum_{n'=1}^{N_j-n} \|\mathbf{r}_j((n'+n)\Delta t) - \mathbf{r}_j(n'\Delta t)\|^2 \quad n = 1,2,...N-1 \tag{3.11}$$

where $\mathbf{r}_j(n'\Delta t)$ is the 2D location of the j-th particle at time $n'\Delta t$, and N_j is the length of the j-th trajectory in frames.

For a specific motion type, the MSD curve fitted through the data points should theoretically follow one of these models described in Table.3.6, where D is the diffusion coefficient, the anomalous exponent ($\alpha < 1$), v is the velocity, r_c is the radius of the confinement and and A_1, A_2 the shape constants. Then the shape of their MSD underlies the motion dynamics and captures the specific diffusion patterns displayed by particles.

Diffusion Type	**Mathematical Description**
Normal Diffusion (ND)	$MSD(n\Delta t) = 4Dn\Delta t$
Directed Motion with diffusion (DM)	$MSD(n\Delta t) = 4Dn\Delta t + (vn\Delta t)^2$
Confined Diffusion (CD)	$MSD(n\Delta t) \simeq MSD(n\Delta t)_c[1 - A_1 exp(-4A_2 Dn\Delta t/MSD(n\Delta t)_c)]$
Anomalous Diffusion (AD)	$MSD(n\Delta t) = 4D(n\Delta t)^{\alpha}$

Table 3.6 Motion Models characterised by the shape of their MSD curve.

The anomalous exponent (α) is the exponent of the model given by $AnomalousDifussion$ (Table.3.6) fitted to the MSD values estimate by Eq.3.11 by power law. It shows values $\alpha \approx 1$ for ND (Brownian motion). However, trajectories showing subdiffusion ($\alpha < 1$) indicate CD ($0 < \alpha < 0.6$) or AD, while those having superdiffusion behaviour ($\alpha > 1$) indicate DM ($\alpha > 1.1$). On the other hand, the moment scaling spectrum (MSS) [193] and its slope (S_{MSS}) was proposed as an approach to improve the calculation of MSD for non-linear diffusion. For each trajectory j the moments of displacement $(\mu_{j,v})$ were calculated for $v = 1,...,6$ as a function of time according to:

$$\mu_{j,v}(n\Delta t) = \frac{1}{N_j - n} \sum_{n'=0}^{N_j-n-1} \|\mathbf{r}_j((n'+n)\Delta t) - \mathbf{r}_j(n'\Delta t)\|^v \tag{3.12}$$

The MSS is just a special case of MSD with $v = 2$. In our implementation, we calculate all moments from $v = 1$ to $v = 6$ for each trajectory by plotting $(\mu_{j,v})$ against $n\Delta t$ in a double logarithmic plot, getting the scaling moments $\gamma_{j,v}$ from assuming each moment μ depends on the time shift according to $\mu_v(n\Delta) \sim n\Delta t^{\gamma\mu}$ [194, 193]. Therefore plotting γ_v against v gives the moment scaling spectrum (MSS) and its slope (S_{MSS}) from linear regression discloses the type of motion [195]: free (S_{MSS} = 0.5), directed (S_{MSS}>0.5), immobile (S_{MSS} < 0.5).

> The method presented in Paper *II* (detailed in *Appendix B*) provides a toolbox for holistic single particle tracking and further user-customizable analysis of tracks.

3.3 Exploring Automated Solutions for BioImage Analysis Pipelines

Automation in bioimage analysis uses computational algorithms, workflows and tools to streamline and simplify the analysis of images. It automates repetitive tasks, reducing manual labor, enhancing efficiency and improving reproducibility. Hence automation reduces human error, enables high-throughput analysis of large datasets and provides valuable information from complex biological images in a efficiently reproducible manner. Biologists can focus on interpretation and discoveries while saving time and effort. Workflow automation tools offer batch processing, parallel computing and integration with other software and databases for systematic, automated analysis, resource utilization along with seamless integration with other scientific tools and data management systems.

This section proposes diverse solutions for handling large and multi-dimensional microscopy images. It covers challenges and consequences of real-time processing in bioimage approaches. Additionally, the concept of open source software is discussed, along with common open source platforms for bioimage analysis.

3.3.1 Dealing with Large and Multi-Dimensional Image Datasets

Recent microscopy advancements enable large volumetric data, posing computational challenges due to scalability and storage limits. This can lead to time-consuming processing and inefficiency in terms of memory accessibility, as data going unprocessed [196]. To address this, automatic approaches streamline image processing and enable efficient handling of large-scale data, by combining storage infrastructure, computational resources, efficient algorithms and automation tools. Moreover, implementing pre-processing operations, such as downsampling or compression, reduces size while preserving key features [90], and decreasing computational demands. Lossy compression also reduces storage needs with minimal information loss whereas, batch processing optimizes resource use and efficient analysis of large datasets with multiple image processing steps. Various platforms (e.g.,CellProfiler, KNIME or Fiji), cloud-based workflow management systems and High-Performance are readily available. ELIXIR [197] coordinates national resources for databases, software tools, cloud storage, high performance computing (HPC) and training. Notebooks such as CodeOcean or Jupyter offer cloud computing and HPC access but lack comprehensive workflow management [198]. These platforms automate data processing, analysis and result generation, ensuring reproducibility, reduced manual intervention by increased productivity.

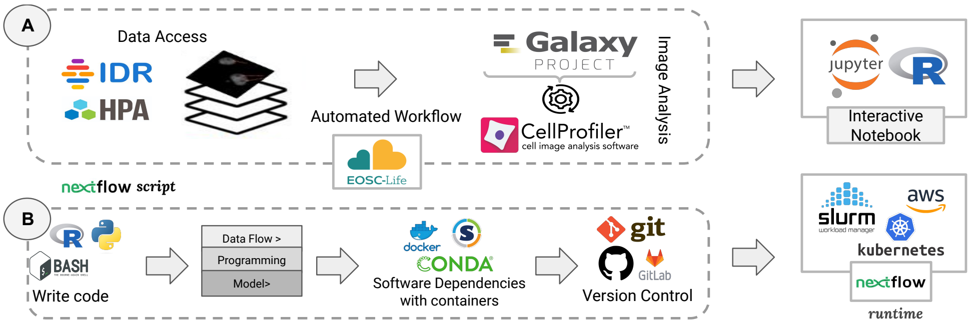

Fig. 3.19 (A) EOSC-Life collaboration integrates CellProfiler into Galaxy Project: Automated workflow by data access (IDR and human protein atlas download), Segmentation and Feature Extraction (19 galaxy tools intregating 22 CellProfiler modules) and Interactive Notebook for downstream Biological analysis (Jupyther and R); (B)

Integrating them into scientific workflow management systems (SWMS) [199] such as Galaxy [200], Nextflow [201], or BIAFLOWS [202] enables comprehensive data integration and execution of complex analysis on large datasets for reproducibility and interoperability with different software. Indeed the integration of CellProfiler modules into Galaxy allows to efficiently handle large datasets in cloud workflows, enabling semi-automated imaging analysis for faster results. This aligns with the EOSC-Life mission, promoting data management, storage, and reuse in the cloud for data-driven research in the life sciences [198]. Additionally, parallel processing [203] can distribute computational workloads across multiple processors or computing resources for efficient handling of computationally intensive tasks on large multi-dimensional datasets. Utilizing parallel computing architectures such as multi-core CPUs and GPUs, or libraries like CLIJ [204], enables creating GPU-accelerated workflows for faster image processing compared to existing acceleration techniques such as ImageJ's batch mode. Furthermore, large multi-dimensional data demand robust storage infrastructure for efficient handling of data volume, retrieval and access. Cloud platforms such as Amazon Web Services (AWS) or Google Cloud Platform (GCP) provide scalable and on-demand computing resources, enabling flexible storage and processing without significant upfront infrastructure investment. Scalable solutions such as network-attached storage (NAS) enable seamless data transfer from microscope-connected computers to analysis workstations, eliminating the need for external hard drives or remote cloud-based storage.

3.3.2 Real-Time Processing for BioImage Analysis

Real-time processing is an emerging field with numerous applications [205]. It plays a crucial role in rapid decision-making scenarios, such as tumor identification [206–209], traffic monitoring [210, 211], facial recognition, assessing plant health [212, 213], or remote sensing [214–216]. Real-time applications continuously interact with the environment they are designed to control. These applications receive input from it, processing input, reacting to changes, and generating appropriate outputs or altering their internal state [205]. Within bioimage analysis, it implies processing images in real-time or near it, typically completing the processing before acquiring the next image. In microscopy, it can be referred to process images on-the-fly as they are being acquired by the microscope in near-real-time, without storing the dataset or causing time processing delays. This approach enables rapid feedback, critical for live cell imaging and high-throughput screening, reducing human influence and saving time for post-processing. Real-time techniques have been applied in cryo-EM [217, 218] for movie alignment, CTF stimation and particle picking. In optical microscopy [219], used by including ML supervised algorithm for cell counting and label-free classification. Another study [220] presents a self-supervised DL-based model for real-time denoising on a two-photon microscope, achieving high-sensitivity fluorescence imaging. Real-time image processing requires specialized hardware and software to handle large data volumes from current microscopy systems. Specialized hardware, such as GPUs, can parallelize tasks, rapidly processing images as they are being acquired. Streaming processing applies to various bioimage analysis tasks, including registration, segmentation, tracking and feature extraction. It provides real-time feedback during live imaging experiments, such as monitoring fluorescence changes over time. In this context, main challenge is balancing processing speed and accuracy, often necessitating simplified models which could be rapidly computed.

> The method presented in Paper *III* (detailed in *Appendix C*) provides a tool to compensate geometric distortions by performing image registration based on elastic deformations represented by B-splines on-the-fly, while microscope is imaging.

3.3.3 Conceiving the Open Source Software

Open-source software (OSS) emerged from early computing such as TeX typesetting system [221] and GNU operating system [222]. The Open Source Initiative (OSI) formalized it in the late 1990*s* [223], focusing on security, affordability and transparency [224]. Accordingly, the integration of OSS into bioimage analysis has revolutionize the field by offering freely accessible and customizable tools. OSS have democratized bioimage analysis, breaking down financial and institutional barriers and enabling scientists worldwide to engage in advanced image analysis. The open-source encourages knowledge exchange and collaboration within community, regardless of their expertise or location, everyone can contribute to software development. Furthermore, the open-source promotes transparency and reproducibility, as the availability of source code allows others to replicate analyses, accelerating scientific progress by enabling the validation of methods.

> All of the tools presented in Paper I, II, II (detailed in $Appendix\ A, B, C$) are OSS. The source code lives on GitHub.

3.3.4 Common Open Source Software for BioImage Analysis

ImageJ, succeeding NIH Image [225], is a popular OS bioimage platform with 30+ years of development. Its community-driven approach shapes functionalities and bug fixes through contributed plugins. ImageJ's Java runtime adoption expanded its user community. Scriptable nature and Macro language enable automation, even for non-programmers. Additional scripting languages (Groovy, JavaScript, Python...) are now available. ImageJ offers updated docs, code access, a mailing list, and discussion forum for user interaction, sharing workflows and best practices. ImageJ has limitations in handling 5D and struggles with large datasets from advanced imaging modalities. ImageJ2 [226] and SCIFIO [227] address these by expanding functionalities for larger datasets and dimensions. Fiji, inspired by ImageJ, is actively maintained, offering "Update Sites" for plugin management. Another widely used software is CellProfiler suite, including CellProfiler [228] and CellProfiler Analyst [229]. It constructs workflows using "modules" within a "pipeline," with shared $*.cpproj$ files. It suits screening assays and adapts to various imaging experiments, while CellProfiler Analyst blends CellProfiler's measurements with ML-based single cell classification and data visualization. Other software iterations include Bio7 [230], SalsaJ [231], and AstroImageJ [232].New solutions leverage R's power [233] for data analysis, visualization, and management. User-friendly ICY [234] offers diverse plugins and features, with community ratings aiding plugin selection. It integrates ImageJ, enabling image and ROI exchange. Its GUI enables to create batch-ready "protocols" shared as scripts in JavaScript or Python. Ilastik [154] supports ML image segmentation, annotation, pixel/object classification, and tracking as well as interoperability with other softwares. Moreover, QuPath [235] for digital pathology handles large images ($> 50kx50kpixels$), integrates with ImageJ, and supports Groovy scripting (own QuPath API). Finally, Bio-Formats [79] ensures file format interoperability among software. KNIME [236] is a workflow createor which integrates APIs of varios software through a GUI for seamless workflow integration. See Table.3.7 for further details on open-source tools described above.

URL	cellprofiler.org	fiji.sc	imagej.net	icy.bioimageanalysis.our	qupath.github.io	knime.com
Developer	Broad Institute of Massachusetts, Institute of Technology	National Institutes of Health and Laboratory for Optical and Computational Instrumentation		Institut Pasteur and France-Bioimaging	Northern Ireland Molecular Pathology Laboratory, Centre for Cancer Research and Cell Biology	University of Konstanz, Zurich, Switzerland
Usage	GUI	GUI		GUI	GUI	GUI
Functionalities	Basic Image processing functionalities and image measurements	Image and video processing functionalities		Visualization, annotation and quantification of bioimaging data	Image Processing for digital pathology and whole slide image	Modular environment, which enables easy visual assembly and interactive execution of a data pipeline
Image Analysis	Manual	Semi-automatic		Semi-automatic	Semi-automatic	Semi-automatic
Reference	McQuin et al.,2018	Schindelin et al.,2012	Schneider et al.,2012	Chaumont et al.,2012	Bankhead et al.,2017	Berthold et al.,2009

Table 3.7 Table listing open source and licensed software tools for bioimage analysis.

3.3.5 Deep Learning Open-source Tools for BioImage Analysis

The bioimage analysis community develops user-friendly OSS as described in Section.3.3.4 . ML-based software such as Weka [133] or Ilastik [237] offer user-friendly solutions. Recently, these platforms have integrated DL-based approaches, and new tools have emerged to make it accessible to non-programmers (detailed in Table.3.8). Using pre-trained DL models involves making predictions on new data without training or parameter tuning. CellProfiler and Ilastik offer pre-trained U-net models for various image analysis tasks along with model training (existing GT or from scratch), well-documented and support. ImageJ, Fiji, and Napari have plugins for pre-trained models such as U-net [77], StarDist [138], and Cellpose [141]. DeepImageJ [238] enables the user-friendly integration of DL models within ImageJ. The Bioimage Model Zoo [239] is a community-driven repository centralizing and promoting the reuse of published DL models in bioimage analysis, expected to become a reference model's resource in this field. On the other hand, manually producing high-quality ground-truth annotations for training can be tedious, especially for 3D+time datasets. Tools such as CellPose or AnnotatorJ [240] have eased the annotation process for 2D and 3D datasets. ImJoy is a web-based platform providing interactive GUI for ground-truth annotation on multi-dimensional images, pre-trained models, and model training. ZeroCostDL4Mic [188] is a Google Colab Python toolbox for DL model training or prediction with no programming knowledge. CSBDeep toolbox is a well-maintained resource, operable from Python or Fiji, providing extensive documentation and facilitating the reuse of DL models (denoising, restoration, and segmentation).

					ZeroCostDl4Mic				CSBDeep
URL	**cellprofiler.org**	**ilastik.org**	**deepimagej .github.io**	**imjoy.io**	**github.com/HenriquesLab/ZeroCostDL4Mic**	**bioimage.io**	**napari.org**	**cellpose.org**	**csbdeep. bioimagecomputing.com**
Type	GUI-based	GUI-based	ImageJ/Fiji plugin to predict on pretrained models	Online computing platform for DL bioimage analysis pipelines	Google Colab Python notebooks implementing DL algorithms	Community driven online repository for DL models	Napari plugins for DL training and prediction on pretrained models	Python DL toolbox for training and prediction	ImageJ/Fiji and Python DL toolbox for general bioimage analysis
Use-Case	Inference with pre-trained models and model training	Inference with pre-trained models and model training	Inference with pre-trained models	Inference with pre-trained models and model training	Inference with pre-trained models and model training	Retrieve models architecture, and pre-trained weights	Inference with pre-trained models and model training	Inference with pre-trained models and model training	Model training from existing ground truth for image restoration
Requirements	None	None	Experience with ImageJ/Fiji	None	None	Dependent on the pretrained model	None	None	Experience with Fiji/ImageJ or Python
Reference	**McQuin et al.,2018**	**Berg et al.,2019**	**Gómez-de-Mariscal et al.,2019**	**Ouyang et al.,2019**	**von Chamier al.,2021**	**Ouyang W et al.,2021**	**Chiu C et al., 2022**	**Stringer C et al., 2021**	**Not Applicable**

Table 3.8 Table listing open source and licensed software tools for bioimage analysis based on deep learning.

Chapter 4

Methodology, Contributions and Applications per Paper

Biological processes are influenced by numerous factors which are often only observable under specific conditions, requiring various imaging techniques to capture them. Integrating this information through image analysis pipelines necessitates the development of advanced image processing tools. Thus the development of these tools for bioimage analysis has revolutionized the field of biological research, enabling quantitative analysis and extraction of valuable information from complex microscopy acquisitions. This thesis presents a diverse range of approaches which bridge the gap among computer science and biology, facilitating knowledge exchange between both disciplines. The primary accomplishments of this multidisciplinary thesis center around the development of user-friendly image processing tools, achieved through the implementation of advanced techniques for microscopy images. These cutting-edge methods focus on enhancing the accuracy and efficiency of bioimage analysis tasks, by the implementation of automation, batch analysis and streaming processing. These methods encompass tasks such as image enhancement, segmentation, feature extraction, cell-type classification, single particle tracking, image registration and visualization. Therefore, leveraging the power of computer vision algorithms, these developed methods enable automated analysis, yielding more reliable and reproducible analyses than manual approaches.

The tools presented in this thesis make significant contributions to the field by proposing easy-going tools tailored for daily bioimage analysis tasks. These contributions aim to address the challenges of processing data from large-scale, low-resolution and multidimensional microscopy datasets. These tools find applications in diverse areas of biological research such as cellular structure analysis, dynamic process tracking, molecular interaction quantification, subcellular localization studies and the characterization of complex biological systems. In this regard, their potential usability might be extended to additional fields such as neuroscience, developmental biology, pathology, and drug discovery, supporting various imaging modalities, including fluorescence microscopy, electron microscopy, correlative microscopy or super-resolution microscopy. Additionally, through the following listed papers, we demonstrate the application of them in real-world bioimage datasets, yielding results which prove their effectiveness in analyzing and extracting relevant information from complex biological images. The evaluation metrics, performance comparisons, and case studies showcase advantages and capabilities of the purposed tools, underscoring their potential impact on advancing biological research.

This chapter serves as a comprehensive compilation of the contributions, methods, applications and results achieved in this thesis. The presented scientific papers exemplify the field advancements, providing effective solutions for daily image processing routines in various microscopy facilities. First, we discuss our contribution to semi-automated cell-type classification by presenting `Cell-TypeAnalyzer` plugin (*Paper I*). Then we discuss our contribution to semi-automated single-particle tracking, diffusion/intensity analysis and subsequent track motion classification by introducing `TrackAnalyzer` plugin (*Paper II*). Finally, we present our contribution to real-time correction of geometrical distortions using a B-spline based elastic registration technique by proposing `OFM-Corrector` protocol (*Paper III*).

4.1 Paper I: Cell-TypeAnalyzer: A flexible Fiji/ImageJ plugin to classify cells according to user-defined criteria

Currently, fluorescent imaging and labeling techniques are commonly used to identify important biological processes by extracting quantitative data from labeled molecules of interest. Advances in open-source software and scientific computing, as well as automation and analysis algorithms, have improved reproducibility and objectivity in cell counting and single-particle analysis, reducing the need for manual analysis. However, classifying specific cell types based on morphology or phenotype remains a labor-intensive and subjective task. Designing a versatile algorithm to automatically identify different cell types on multiple fluorescent markers is challenging, especially considering low signal-to-noise ratios and limited resolution in fluorescence microscopy. To address these challenges, `Cell-TypeAnalyzer` is an open-source plugin which enables the classification of cells based on morphological, intensity, or spatial features (detailed in Fig.4.1 (A-B)). Although the concept of `Cell-TypeAnalyzer` for cell-type classification is not new, the contribution of this paper is to offer a semi-automated cell-type analysis, instead of fully automated mode, potentially compromising accuracy and user interpretation. `Cell-TypeAnalyzer` allows researchers to describe a cell population through a set of extracted features, identifying biologically relevant similarities or variations. The tool is highly configurable and can be adapted to various imaging conditions by manually adjusting internal parameters. Reaching an accurate segmentation is crucial within `Cell-TypeAnalyzer` workflow, particularly when dealing with heterogeneous background or touching cells since the analysis is carried out for each cell. For such, `Cell-TypeAnalyzer` integrates from MorpholibJ library, auto-threshold global methods to automatically segment images [126], without assuming binary shapes or circularity, as well as algorithms for watershed segmentation and morphologial operators. After segmentation, each cell is individually measured and described using physical, geometrical, morphological, statistical, and intensity-based features for further user-customized cell-type classification.

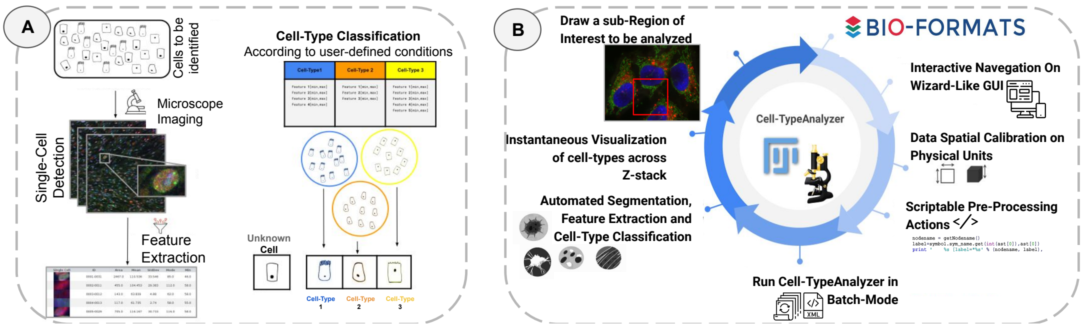

Fig. 4.1 Different aspects of `Cell-TypeAnalyzer` plugin. (A) Illustration of the workflow to identify specific cell-types in a cell population: (B) Schematic description of `Cell-TypeAnalyzer` main functionalities

`Cell-TypeAnalyzer` was developed within the ImageJ ecosystem, benefiting from the platform's image processing capabilities. It offers a wizard-like GUI which guides researchers through each step of the analysis, providing instant visualization of outputs for each marker and allowing manual, visual and quantitative verification. The plugin can process large sets of images, supporting multiple image formats (Bio-Formats library) and allowing users to define specific regions of interest to be analyzed. While other tools widely used for cell-type analysis operate in fully automated or unsupervised mode, `Cell-TypeAnalyzer` functions in semiautomatic mode, require user input and interaction for accurate feature extraction. It addresses the need for universal tools for semiautomated cell-type analysis within the ImageJ or Fiji ecosystem, offering a solution which relies less on full automation and thus is more reliable given the current limitations of fully automated methods. It also provides options for customizing cell-type analysis on multi-fluorescent microscopy images and can be applied to various microscopy samples. Compared to similar software for cell-type analysis, `Cell-TypeAnalyzer` allows users to tune classification parameters and provides a simple GUI for visualizing and verifying outputs. It leverages existing ImageJ plugins and libraries, offering a wide range of tools for image preprocessing and analysis. `Cell-TypeAnalyzer` does not remove noise or enhance image quality but provides scriptable (ImageJ's Macro Language) functionality and by default pre-processing tools (denoising, filtering and contrast enhancer) to improve cell detection and characterization. Best practices for image acquisition are recommended before using the tool.

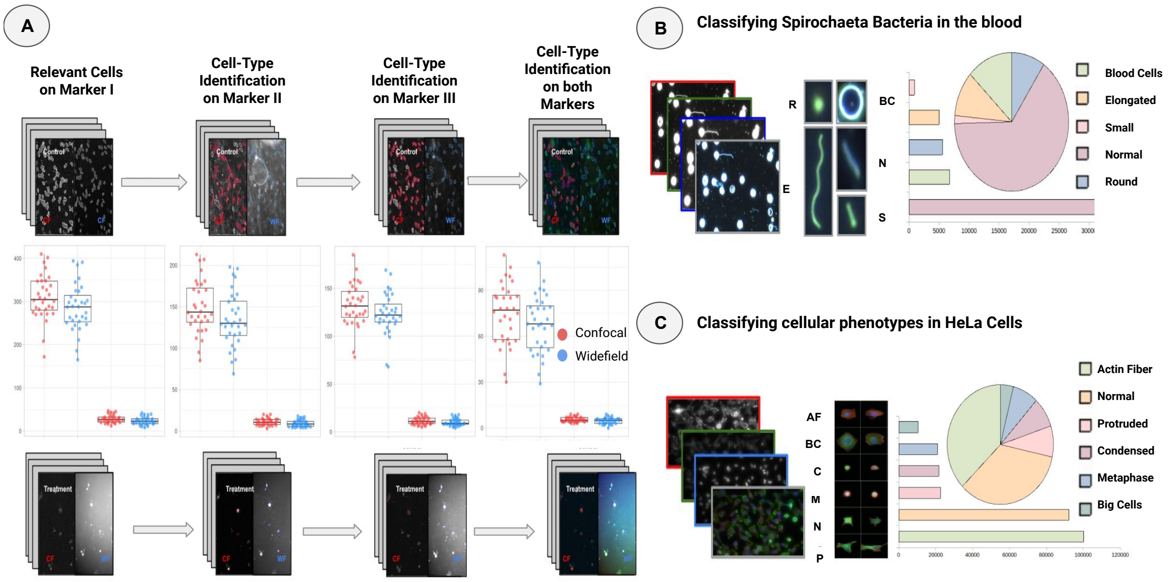

Fig. 4.2 Experimental Validation by different applications using different datasets. (A) Quantification count of cells of a given type using Confocal or Widefield microscopy, showing that imaging with both modalities does not make any significant difference for this experiment (at a confidence level of 95 %); (B) Semi-automated analysis for classifying cellular phenotypes in HeLa cells into Actin Fiber (AF), Big cells (BC), Condensed (C), Metaphase (M), Normal (N), and Protruded (P) ;(C) Semi-automated analysis for classifying Spirochaeta bacteria in the blood into Blood Cells(BC), Round(R), Elongated(E), Small(S) and Normal(N).

Furthermore, `Cell-TypeAnalyzer` supports batch processing, making it suitable for analyzing large microcopy datasets. It provides data visualization options, including dynamic scatter plots, to explore and drill down into cell-type classification results. The tool is designed to be user-friendly and does not require programming proficiency. It is open-source and freely available, with documentation and video tutorials provided on its GitHub page `https://github.com/acayuelalopez/CellTypeAnalyzer`. On the other hand, we showed in our experimental validation (shown in Fig.4.2) that our tool successfully compares the count of cells of a specific type using both Confocal and Widefield microscopy techniques. The results indicate that there is no statistically significant difference among the two imaging mode for this experiment, with a confidence level of 95% (Fig.4.2 (A)). Moreover, we achieved successful classification of cellular phenotypes in HeLa cells (Fig.4.2 (C)) based on morphological changes, with the proportions of cells in each type being similar to those originally reported in [241]. Additionally, we conducted morphological phenotyping of Spirochaeta bacteria in blood (Fig.4.2 (B)).

In summary, `Cell-TypeAnalyzer` combines semiautomatic, scriptable and manual tools to achieve accurate cell-type analysis, even in images with overlapping objects. It offers efficient batch processing and allows the identification of hundreds of cell types per minute. Despite some limitations, `Cell-TypeAnalyzer` is accessible to researchers at different levels of expertise, facilitating cell-type classification under user-defined conditions.

4.1.1 Publication Summary

Further details about the proposed solutions toward cell-type classification using Cell-TypeAnalyzer can be found in the following relevant authored publication:

> Cayuela López, A., Gómez-Pedrero, J., Blanco, A., & Sorzano, C. (2022). Cell-TypeAnalyzer: A flexible Fiji/ImageJ plugin to classify cells according to user-defined criteria. *Biological Imaging*, 2, E5. doi:10.1017/S2633903X22000058 [242]

The complete publication is enclosed as *Appendix A*.

4.2 Paper II: TrackAnalyzer: A Fiji/ImageJ Toolbox for a holistic Analysis of Tracks

The advancement of innovative imaging techniques, particularly Total Internal Reflection Microscopy (TIRF), has become vital for studying dynamic processes within cells at the sub-cellular level. These techniques enable quantitative analysis of intracellular dynamics with high spatial resolution (tens of nanometers) and long-term observation capabilities. Single-particle tracking analysis has emerged as a standard tool in life sciences, facilitated by fluorescent protein labeling and software advancements. SPT allows real-time measurement of motion, diffusion properties and spatial distribution changes of single particles with high temporal resolution and signal-to-noise ratio.

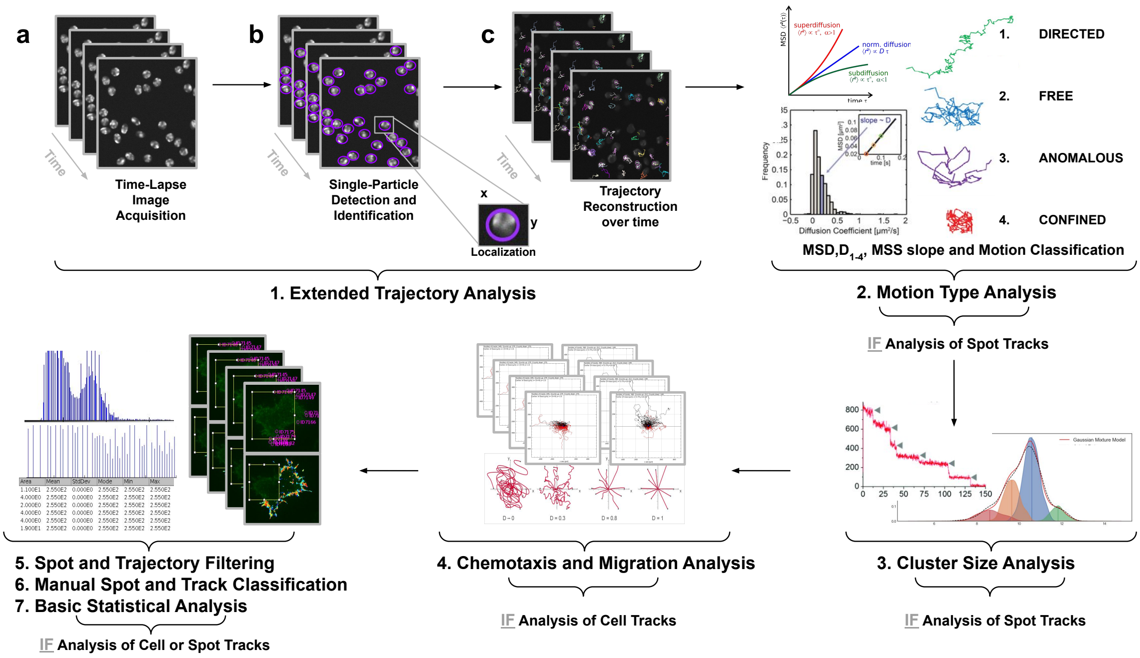

Fig. 4.3 Illustration of the workflow to perform single particle tracking together with subsequent analysis of diffusion using `TrackAnalyzer` software which consists of several processes.

SPT algorithms address various challenges inherent to cellular dynamics, such as gap-closing, merging and splitting events, to achieve accurate tracking of fluorescent particles. These algorithms aim to reconstruct the motion of particles over consecutive time points by bridging missing detections and tracking their movements. To ensure reliable tracking, high signal-to-noise ratios are crucial, as noise from background fluctuations, autofluorescence, blinking, photobleaching, phototoxicity, poor contrast, high particle density, and motion heterogeneity can affect the accuracy and reproducibility of the analysis. In this regard, SPT analysis involves both spatial (particle detection) and temporal (particle linking) methods. Spatial methods segment and locate each spot or cell, establishing frame-by-frame correspondences in X-Y-T coordinates. Conversely, temporal methods assign detected single particles to individual tracks over time.

While manual tracking is feasible for low particle densities, automation is preferred due to its objectivity, reproducibility, and efficiency, specially ofr high particle densities. However, fully automated SPT approaches may encounter challenges with changes in experimental conditions, making a combination of automation and user control desirable for temporal quantitative analysis. Existing software tools for SPT analysis lack user-friendliness and comprehensive functionality to handle diverse time-lapse microscopy acquisitions. To address this gap, `TrackAnalyzer` software was developed. `TrackAnalyzer` allows users to set up customized SPT analyses based on their own experimental conditions and apply them in batch-mode to multiple time-lapse acquisitions. The software, available as an open-source plugin for Fiji or ImageJ, offers functionalities for spot detection, track reconstruction, diffusion analysis, trajectory analysis, cluster size analysis, single-step photobleaching analysis, and integration with the Chemotaxis and Migration Tool for quantifying chemotaxis and migration experiments. `TrackAnalyzer` introduces three key contributions to enhance SPT analysis. Firstly, it offers an user-friendly wizard-like GUI, that enables batch-mode analysis, allowing users to automatically analyze multiple datasets by configuring the analysis for one dataset and replicating it across others within the same experiment. This addresses the limitation of main existing tools which only allow analysis of a single time-lapse dataset. Secondly, `TrackAnalyze` leverages TrackMate, an open-source software, to provide flexible and adaptable algorithms for spot detection and track reconstruction over time; TraJClassifier integration to locally and globally both characterize and classify trajectory motion into normal diffusion, subdiffusion, confined diffusion and directed/active motion by a random forest approach. Chemotaxis and Migration Tool integration which allows advanced analysis of chemotaxis experiments. Therefore, `TrackAnalyzer` implements additional features such as cluster size, intesity analysis, motion analysis and track classification. Third, it benefits from the extensive ImageJ ecosystem, integrating various plugins for scientific image processing. By combining these functionalities with existing powerful tools, `TrackAnalyzer` offers comprehensive suite of tools which facilitate analysis of particle behavior under diverse experimental conditions, allowing for quantitative comparisons of particle parameters. On the other hand, we showed in our experimental validation (shown in the Fig.4.4) that our tool successfully does the analysis of the dynamic of CXCR4 at the plasma membrane of Jurkat CXCR4$^{-/-}$ cells electroporated with CXCR4-AcGFPm (detailed in Fig.4.4(A)), according to the results originally reported in [243]. Additionally, we conducted the analysis of the directed cell migration capacity of Jurkat cells to illustrate TrackMate features to evaluate directional cell migration by using Chemotaxis and Migration Tool (detailed in Fig.4.4(B)).

In conclusion, `TrackAnalyzer` significantly enhances the capabilities of SPT analysis by providing a wizard-like GUI, batch-mode analysis, extended analysis functionalities and integration with existing software tools. It empowers researchers in the quantitative analysis of particle behavior under various experimental conditions, contributing to a deeper understanding of dynamic processes within cells at the sub-cellular level.

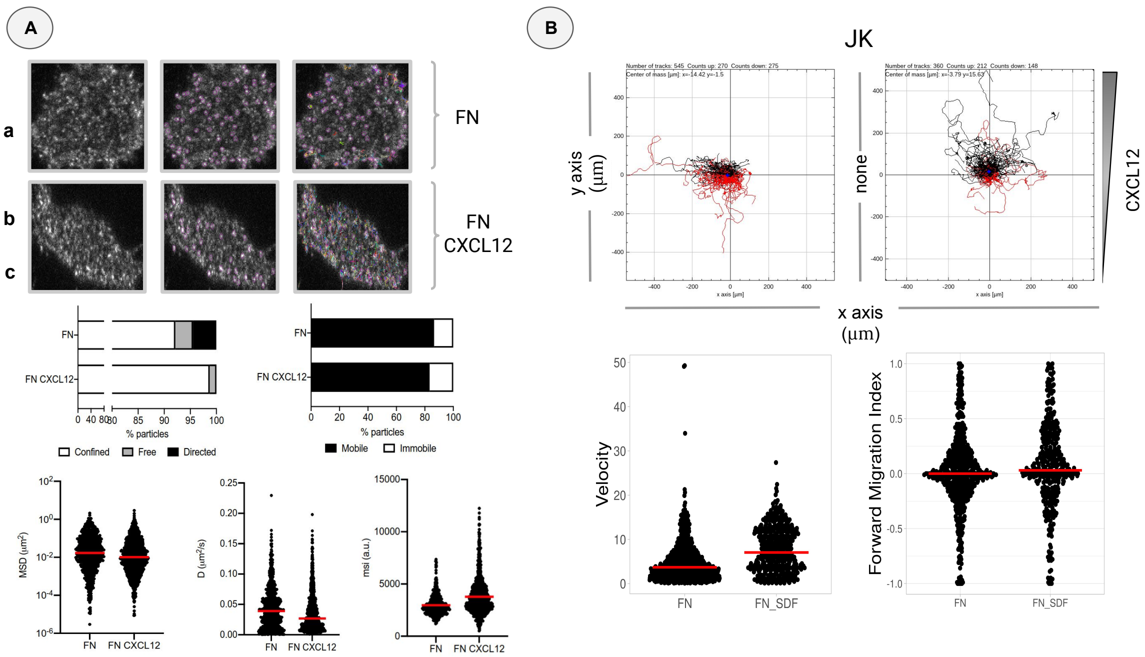

Fig. 4.4 Example of Experimental Validation by different applications using different datasets. (A) Application of `TrackAnalyzer` to track CXCR4-AcGFPm in JK CXCR4$^{-/-}$ cells electroporated with CXCR4-AcGFPm. (B) Migration of JK cells in response to a CXCL12 gradient.

4.2.1 Publication Summary

Further details about the proposed solutions toward single particle tracking analysis and subsequent track analysis using `TrackAnalyzer` can be found in the following relevant authored publication:

> Cayuela López, A., García-Cuesta, E., Gardeta, S., Rodríguez-Frade, J., Mellado, M., Gómez-Pedrero, J., & S. Sorzano, C. (2023). TrackAnalyzer: A Fiji/ImageJ Toolbox for a holistic Analysis of Tracks. *Biological Imaging*, 1-14. doi:10.1017/S2633903X23000181 [244]

The complete publication is enclosed as *Appendix B*.

4.3 Paper III: Real-Time Correction of Chromatic Aberration in Optical Fluorescence Microscopy

In recent years, significant advancements have been made in single molecule-based super-resolution microscopy techniques. Multi-color fluorescence imaging stands out among these techniques, enabling the differentiation of proteins and structures of interest in both living and fixed cells. However, challenges such as mechanical drift and chromatic aberrations still remain decreasing the image resolution. While chromatic aberration is a common problem in multi-color imaging, other factors such as imperfect optical elements, refractive index mismatches, and dispersion in biological samples can lead to geometric distortions.

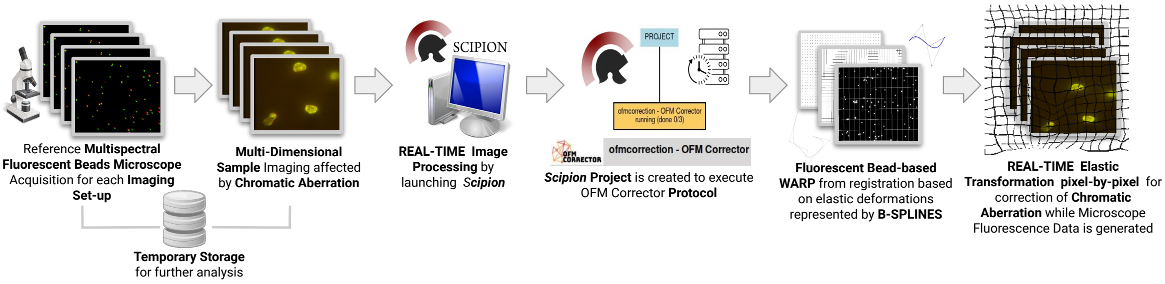

Fig. 4.5 Graphical representation of the workflow to reach real-time correction of geometric misalignment among channels in multi-dimensional images acquired with fluorescence microscopy using multi-spectral fluorescent beads through Scipion software.

Total Internal Reflection Fluorescence (TIRF) microscopy emerges as a potent technique for selectively imaging molecules in an aqueous environment with a high refractive index. Its thin axial optical sectioning and high signal-to-noise ratio make it suitable for imaging membrane-associated events in living cells and molecules at the medium interface. Nevertheless, TIRF microscopy is susceptible to lateral chromatic aberration, which introduces shifts, rotations and scaling differences among color channels. To address chromatic aberration in TIRF microscopy, elastic (non-rigid) image registration techniques are employed, involving the alignment of corresponding features in multiple images through a geometric transformation. B-spline-based elastic image registration is implemented to handle a wide range of deformations, including non-linear ones. This method ensures high-quality interpolation and localized control over the deformation field, providing a practical solution which surpasses merely improving the physical construction of the microscope's dichroic mirror. This B-spline-based elastic image registration method is integrated into the `OFM Corrector` protocol, freely available within the Scipion framework. The Scipion framework allows for real-time or stream processing for image registration, enabling almost instant aberration-corrected images on-the-fly while microscope is imaging. The protocol offers a unified graphical user interface, package interoperability and workflow monitoring for streaming elastic image registration. The software-based approach compares favorably to expensive optical solutions and is more versatile, addressing not only chromatic aberration but also other sources of geometric distortions.

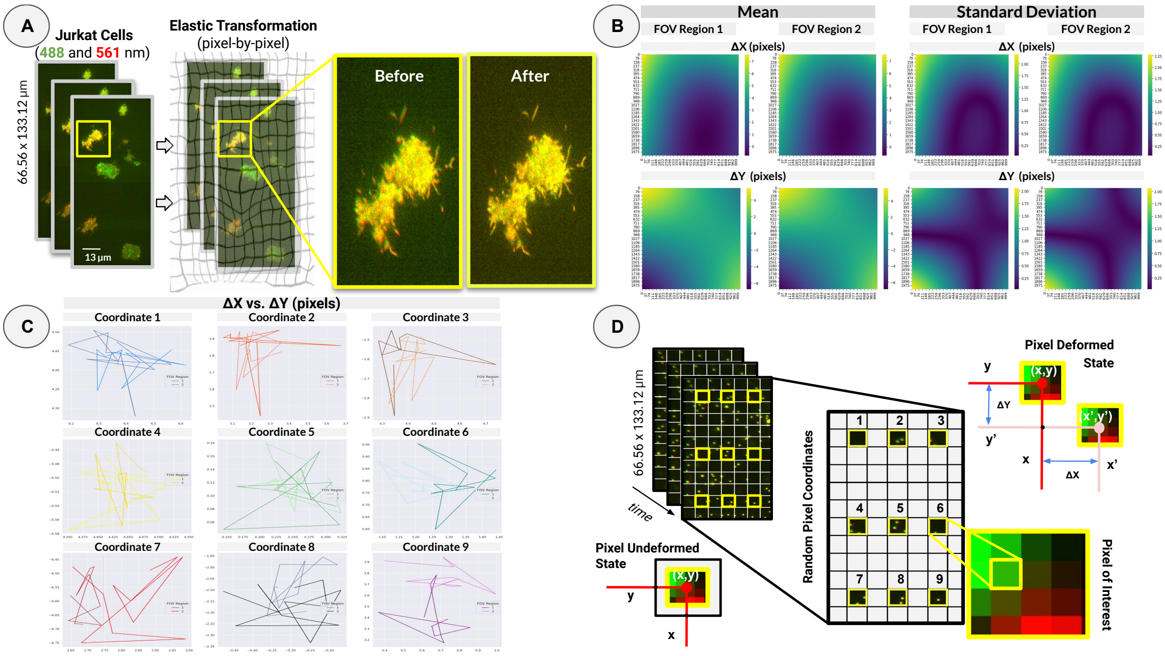

Fig. 4.6 (A) The deformation field is corrected by using `OFM-Corrector` for all the input videos corresponding to biological samples; (B) Mean and standard deviation of the deformation field over time at two different regions of the FOV. The size of the acquired region is 1024x2050 pixels (66.56x133.12 μm); (C) For some representative coordinates, we show the trajectory over time of the deformation field. Displacements are expressed in pixels (the pixel size is 0.065 μm; (D) Schematic visualization of the shift measurement procedure. x,y are the pixel coordinates (undeformed state) and x',y' (deformed state) are the pixel coordinates after elastic transformation, and ΔX, ΔY are the displacement among them.

In this pipeline (detailed in Fig.4.5), multispectral fluorescent beads are used as a reference for elastic image registration and drift correction. These beads emit differently in the same wavelength range as the applied dyes, allowing for the registration of chromatic shifts. This approach can be easily applied to daily microscopy routines in facilities by capturing a reference calibration image for each specific imaging setup, considering factors such as excitation laser lines, objective lens, temperature stability, and exposure time. It is essential to select suitable multi-spectral fluorescent beads based on their signal and size to ensure they exceed the microscope's resolution, providing a sufficient signal-to-noise ratio. The software solution is not limited to correct geometrical distortions among two channels but can simultaneously correct any number of channels. One channel is chosen as the reference, and the others (target channels) are corrected to match the reference. In addition, we experimentally compute the mean and standard deviation of the deformation field over time at two different regions of the FOV (shown in Fig.4.5(B)). We also evaluate the stability of the deformation field over the time for some representative coordinates (shown in Fig.4.5(C)).

In conclusion, this `OFM-Corrector` contributes as an efficient and cost-effective approach to correct chromatic aberration deformations and other sources of geometric distortions. It can be seamlessly integrated into standard procedures in microscopy facilities and is not limited to specific imaging setups. Future work could focus on expanding the protocol capabilities to address additional optical distortions commonly encountered in light imaging.

4.3.1 Publication Summary

Further details about the proposed solutions toward real-time correction of chromatic aberration by B-spline-based elastic image registration analysis using `OFM-Corrector` can be found in the following relevant authored publication:

> López AC, Conesa P, Oña Blanco AM, Gómez-Pedrero JA, Sorzano COS. Real-time correction of chromatic aberration in optical fluorescence microscopy. *Methods Appl Fluoresc*. 2023 Jul 3;11(4). doi: $10.1088/2050-6120/ace153$. PMID: 37352866.[245]

The complete publication is enclosed as *Appendix C*.

Chapter 5

List of Publications

5.1 Publications used for the compendium of articles

Below are listed (co-)authored publications used for this dissertation, structured as a compendium of articles, in chronological order.

> Cayuela López, A., Gómez-Pedrero, J., Blanco, A., & Sorzano, C. (2022). Cell-TypeAnalyzer: A flexible Fiji/ImageJ plugin to classify cells according to user-defined criteria. *Biological Imaging*, 2, E5. doi:10.1017/$S2633903X22000058$ [242]

- In this paper, we present our semiautomated, wizard-like GUI and versatile tool, which aims to eliminate the time-consuming and biased manual cell-type classification. This powerful tool functions as an open-source plugin for Fiji or ImageJ, enabling us to detect and classify cells in 2D images effectively. Our workflow comprises several essential steps, including: (a) Image preprocessing actions, data spatial calibration, and ROI definition; (b) Segmentation; (c) Extraction of cell features; (d) Filters to select relevant cells; (e) Definition of specific criteria to categorize cells into distinct cell types.; (f) Cell-type classification; (g) Flexible analysis of the results to gain meaningful insights. Moreover, `Cell-TypeAnalyzer` supports batch processing. We experimentally show that our tool is able to compare the count of cells of a given type using Confocal or Widefield microscopy, showing that imaging with confocal or widefield microscopy does not make any statisti-cally significant difference for this experiment (at a confidence level of 95%). In addition, we classified cellular phenotypes in HeLa cells based on morphological changes, being the proportions of cells in each one of the types similar to the one originally reported in [241]. Finally, morphological phenotyping of Spirochaeta bacteria in blood.

- I have designed the program, implemented the algorithms and wrote the software code and performed all validation experiments. I am also the main author of the manuscript.

Cayuela López, A., García-Cuesta, E., Gardeta, S., Rodríguez-Frade, J., Mellado, M., Gómez-Pedrero, J., & S. Sorzano, C. (2023). TrackAnalyzer: A Fiji/ImageJ Toolbox for a holistic Analysis of Tracks. *Biological Imaging*, 1-14. doi:10.1017/S2633903X23000181 [244]

- In this paper, we introduce our newly developed tool, `TrackAnalyzer`, accessible from Fiji and ImageJ. This versatile tool facilitates the execution of semi-automated Single-Particle Tracking (SPT) and subsequent motion classification. Additionally, it enables quantitative analysis of diffusion and intensity for selected tracks by handling large sets of time-lapse images. It supports feature extraction and user-defined classification for further analysis. Our analysis workflow is designed to automate the following key steps: (a) Spot detection and filtering; (b) Construction of tracks; (c) Track classification and analysis, including diffusion and chemotaxis assessments. (d) Detailed analysis and visualization of outputs. Our pipeline is semi-automated and it enables batch processing. By providing an accessible solution for live-cell imaging analysis, it contributes to advancing our understanding of biological processes with enhanced accuracy and efficiency. Its user-friendly GUI and versatile functionalities make it a valuable tool for the scientific community. We experimentally show that our tool is able to do the the analysis of the dynamic of CXCR4 at the plasma membrane of Jurkat CXCR4/ cells electroporated with CXCR4-AcGFPm, according to the results originally reported in [243]. Additionally, we conducted the analysis of the directed cell migration capacity of Jurkat cells to illustrate TrackMate features to evaluate directional cell migration by using Chemotaxis and Migration Tool.
- I have designed the program, implemented the algorithms and wrote the software code and performed all validation experiments. I am also the main author of the manuscript.

López AC, Conesa P, Oña Blanco AM, Gómez-Pedrero JA, Sorzano COS. Real-time correction of chromatic aberration in optical fluorescence microscopy. *Methods Appl Fluoresc*. 2023 Jul 3;11(4). doi: $10.1088/2050-6120/ace153$. PMID: 37352866 [245]

- In this paper, we present an innovative extension of Scipion named `OFM- Corrector`, which enables real-time correction of geometrical distortions using a B-spline-based elastic continuous registration technique. Our proposal offers a straightforward approach to compensate chromatic aberration, digitally realigning color channels in multi-color microscopy images, even when dealing with 3D or temporal data. The core of our method involves the utilization of fluorescent beads excited by different wavelengths of light. By registering them, we obtain an elastic warp as a reference for correcting chromatic shifts. Our software is freely and readily available for those working in light microscopy facilities. With the integration of `OFM-Corrector` into Scipion's image processing framework, we aim to empower researchers with improved image resolution and accuracy, fostering deeper insights into the intricate world of sub-cellular interactions and structures. We experimentally compute the mean and standard deviation of the deformation field over time, and we evaluate the stability of the deformation field over the time.

- I have designed the program, implemented the algorithms and wrote the software code and performed all validation experiments. I am also the main author of the manuscript.

5.2 Other publications

Below are listed other (co-)authored publications in chronological order.

> Cuesta-Geijo MÁ, García-Dorival I, Del Puerto A, Urquiza J, Galindo I, Barrado-Gil L, Lasala F, Cayuela A, Sorzano COS, Gil C, Delgado R, Alonso C. New insights into the role of endosomal proteins for African swine fever virus infection. *PLoS Pathog*. 2022 Jan 26;18(1):*e*1009784. doi: 10.1371/*journal.ppat*.1009784. PMID: 35081156; PMCID: *PMC*8820605.

- In this paper, authors evaluate the role of African swine fever virus (ASFV) which infects cells through endocytosis and relies on interactions with endosomal proteins for successful fusion. NPC1 and Lamp-1/-2 play crucial roles in this process. Understanding these interactions could shed light on ASFV infection.
- I have designed the ImageJ plugin, implemented the algorithms and wrote the software code to quantify the number of viral cores trapped within the enlarged Rab7+vesicles.

> Soler Palacios B, Villares R, Lucas P, Rodríguez-Frade JM, Cayuela A, Piccirillo JG, Lombardía M, Delgado Gestoso D, Fernández-García M, Risco C, Barbas C, Corrales F, Sorzano COS, Martínez-Martín N, Conesa JJ, Iborra FJ, Mellado M. Growth

hormone remodels the 3D-structure of the mitochondria of inflammatory macrophages and promotes metabolic reprogramming. *Front Immunol*. 2023 Jul 5;14:1200259. doi: 10.3389/*fimmu*.2023.1200259. PMID: 37475858; PMCID: *PMC*10354525.

- In this publication, authors demonstrate that GH likely serves a modulatory role in the metabolism of inflammatory macrophages and suggest that metabolic reprogramming of macrophages should be considered as a new target to intervene in inflammatory diseases.
- I have designed the ImageJ workflow, implemented the algorithms and wrote the software code to evaluate 3D mitochondrial morphology and network connectivity in fusion-fission events for both fluorescence and cryo-FIBSEM images.

Chapter 6

Open-source Code and Data Availability

Promoting scientific reproducibility relies heavily on the release of open-source code. Furthermore, facilitating the utilization of cutting-edge technologies through user-friendly tools significantly enhances analytical processes. As a result, promoting a culture which prioritizes open-source initiatives becomes imperative for advancing scientific research. Consistent with this principle, the authors aim to provide explicit references to the repositories and online resources developed throughout this thesis, all of which are freely accessible:

Paper I: Cell-TypeAnalyzer: A flexible Fiji/ImageJ plu- gin to classify cells according to user-defined criteria

Cell-TypeAnalyzer plugin:

> The github repo `https://github.com/acayuelalopez/CellTypeAnalyzer` is the entry point for all the open-source material created within `Cell-TypeAnalyzer` plugin.

Paper II: TrackAnalyzer: A Fiji/ImageJ Toolbox for a holistic Analysis of Tracks

TrackAnalyzer plugin:

> The github repo `https://github.com/acayuelalopez/TrackAnalyzer_` is the entry point for all the open-source material created within `TrackAnalyzer` plugin.

Paper III: Real-Time Correction of Chromatic Aberra- tion in Optical Fluorescence Microscopy

OFM Corrector protocol:

> The complete source code of the algorithm integrated in Scipion software is available at `https://github.com/acayuelalopez/bUnwarpJ_code`. Our protocol can be used for real-time processing within the Scipion framwork. You can install it by using the Scipion software following the Scipion's installation guide (`https://scipion-em.github.io/docs/release-3.0.0/docs/scipion-modes/how-to-install.html`).

Chapter 7

Conclusion and Future Work

This dissertation, structured as a compedium of articles, was focused on the development of bioimage analysis tools tailored to advanced optical microscopy. The aim was to optimize the extraction of quantitative information from labeled molecules of interest, enabling the discernment of biological processes. This thesis delved into the advancements within open-source software, scientific computation, automation, image analysis algorithms and real-time processing. These advancements have notably improved reproducibility and objectivity of analyses, diminishing the reliance on manual intervention for users at ALMF at CNB. The core aims of this dissertation were to enhance the automation capabilities, shifting from a qualitative and manual analysis paradigm to a large-scale and quantitative assessment of optical microscopy images, analogous to the progress witnessed in electron microscopy at CNB. These goals aimed to improve efficiency and accuracy while transforming the analysis process into a more objective, data-driven and high-throughput approach. However, this thesis acknowledged that classifying specific cell types based on morphology or phenotype remains a labor-intensive and subjective task. To address this challenge, `Cell-TypeAnalyzer` an open-source Fiji plugin was designed to empower the user-customized classification of specific cell-types based on morphological, intensity, or spatial features. This tool offers a semi-automated approach for cell-type classification, providing more objectivity than qualitative strategies, while facilitating a more streamlined and systematic analysis procedure. `Cell-TypeAnalyzer` allows researchers to describe a cell population by a set of extracted features, thus identifying biologically relevant similarities. Its user-friendly GUI seamlessly guides researchers through each step of the analysis, offering instant visualization for each marker and facilitating manual verification. The plugin supports batch processing, making it suitable for analyzing large image datasets. With robust batch processing capabilities, the plugin can handle the cell-type analysis of extensive image datasets.

This dissertation also recognized the significance of automation for a comprehensive analysis of particle behaviour, while emphasized the user control and interpretation. To address the lack of user-friendliness in existing tools for SPT and subsequent motion analysis, we introduced `TrackAnalyzer` plugin. This open-source software allows users to tailor SPT analyses according to their experimental conditions, enabling their application in batch mode across multiple data acquisitions. It offers spot detection, trajectory reconstruction over time, diffusion analysis, trajectory analysis, cluster size analysis, single-step photobleaching analysis, and seamless integration with both TraJClassifier the Chemotaxis and Migration Tool. By providing a user-friendly interface, facilitating batch-mode analysis and extending SPT analysis from TrackMate, `TrackAnalyzer` substantially elevates the landscape of current SPT analysis. This tool empowers researchers to customize and apply SPT analyses based on their experimental contexts, thereby facilitating quantitative comparisons of particle parameters with enhanced ease. Last but not least, this thesis aimed the integration of real-time image processing within the ALMF at CNB. This involved applying real-time elastic (non-rigid) image registration techniques to compensate for geometric deformations induced by chromatic aberration present in the TIRF microscope at ALMF, which decreases the image resolution. To overcome challenges arising from chromatic aberration, this thesis proposed the adoption of an elastic image registration technique based on B-splines. This method was integrated into the *OFM-Corrector* protocol within the Scipion framework, thereby enabling real-time and continuous correction of geometric distortions on-the-fly while the microscope is imaging. By firstly, using multispectral fluorescent beads as a reference, this software provided an efficient and cost-effective solution to compensate chromatic aberration and other geometric distortions. As a result, this protocol can be seamlessly used in conventional microscopy facility practices, enhancing the accuracy and reliability of standard procedures.

By accomplishing these goals, this dissertation wanted to streamline and automate routine image analysis tasks at microscopy facilities, hence ushering in a new era of efficiency at ALMF. We have proposed diverse tools which specifically tackled the challenges of cell-type classification, SPT analysis and real-time correction of geometric distortions, providing user-friendly GUI, fortified with batch processing capabilities and extended analysis functionalities. These tools enhance the accuracy, reproducibility, and efficiency of bioimage analysis, facilitating quantitative comparisons and advancing our understanding of biological processes at the cellular level. Furthermore, this dissertation played a crucial role in the CNB to become a leading center in quantitative biology and bioimaging by establishing a the Quantitative Image Analysis Unit as the central focus. The strategic shift towards massive and quantitative analysis equips the center to proficiently manage large-scale datasets and extract valuable insights from them. In summary, through the achievement of these objectives, this thesis seeks to usher in a new era of heightened efficiency, precision, and objectivity in optical microscopy analysis, thereby contributing to the mission of the CNB to emerge as a leader center in quantitative biology and bioimaging.

7.1 Future Work

The information from these conclusions can be a starting point for future work.

Deep learning shows promise for addressing challenges in optical microscopy, such as low signal-to-noise ratio and variable imaging conditions. In this regard, the future of bioimage analysis involves leveraging cloud computing, developing tailored deep learning architectures, and utilizing machine learning for feature extraction and classification. Integrating these approaches into unified platforms by providing user-friendly GUIs and scalable computing power is encouraging. Notwithstanding these advancements, challenges remain, ranging from ground truth dataset availability and computational resources to efficiently handling large image sizes and voluminous data to model interoperability. Future work should focus on collaborative efforts to create benchmarking datasets, standardized metrics, and best practices for deploying these solutions in optical microscopy analysis.

Regarding the contributions outlined in this thesis, while `Cell-TypeAnalyzer` is presented as a valuable solution for cell-type classification, there is room for improvement. Future work could focus on expanding the algorithm capabilities to handle more complex and diverse cell types, as well as improving its performance in images with low signal-to-noise ratios and limited resolution. The incorporation of multichannel and multi-dimensional image hyperstacks, hence expanding its applicability beyond RGB images, would facilitate the comprehensive analysis of cellular phenomena across different molecular markers and dyes. Additionally, incorporating machine or deep learning techniques could enable the algorithm to learn and adapt to new cell types, further enhancing its classification accuracy. `Cell-TypeAnalyzer` might facilitate compatibility with widely used pre-trained deep learning models for segmentation, such as Stardist or Cellpose. This synergy holds immense promise in further refining cell-type classification accuracy, effectively bridging the gap between traditional analysis and cutting-edge segmentation techniques. Related to the automation in SPT, although `TrackAnalyzer` offers batch-mode analysis and extended functionalities for SPT analysis, further automation could be explored to minimize user intervention/input and enhance accuracy. Developing algorithms which can adapt to changes in experimental conditions, such as variations in imaging parameters or sample characteristics, would improve the robustness and reproducibility of SPT analysis. Integrating pre-trained deep learning models to automate spot detection and track construction could also be a promising direction for future work. Moreover, it is crucial to consider interoperability with other software tools such as Icy for a comprehensive solution

The B-spline-based elastic image registration method implemented in the `OFM-Corrector` protocol provides an effective solution for correcting geometric distortions caused by chromatic aberration in TIRF microscopy. However, future work could focus on developing more comprehensive protocols which can handle multiple types of geometrical distortions simultaneously would further improve image quality and resolution. Moreover, an exciting avenue for exploration involves the integration in the TIRF microscope of real-time single particle tracking analysis approach (main applicability of this microscopy technique) following geometric aberration compensation within the `OFM-Corrector` protocol. In this way, the synergy among real-time dual-channel alignment and subsequent single particle tracking has the potential to enhance time efficiency for users at the ALMF, particularly during the frequent analysis employing the TIRF microscope.

In the middle-term, it would be beneficial to enhance the integration and user-friendliness of bioimage analysis tools. Seamless integration with other existing microscopy platforms, can streamline the analysis workflow and facilitate data exchange between different software tools. Additionally, improving the user interface and providing intuitive visualization options would enhance the usability of these tools, making them accessible to a broader range of researchers with varying levels of technical expertise. As technology advances, new imaging techniques and algorithms will continue to emerge. Future work should focus on keeping up with these advancements and incorporating them into bioimage analysis tools. Exploring cutting-edge algorithms and methodologies, can help maximize the extraction of quantitative information from optical microscopy images, further enhancing the accuracy, resolution, and interpretation of the acquired data.

References

[1] Jason R. Swedlow. Innovation in biological microscopy: Current status and future directions. *BioEssays*, 34(5):333–340, March 2012.

[2] Noël Bonnet. Some trends in microscope image processing. *Micron*, 35(8):635–653, December 2004.

[3] Jianxin Wang, Min Li, Youping Deng, and Yi Pan. Recent advances in clustering methods for protein interaction networks. *BMC Genomics*, 11(Suppl 3):S10, 2010.

[4] Meghan K. Driscoll and Assaf Zaritsky. Data science in cell imaging. *Journal of Cell Science*, 134(7), April 2021.

[5] Fazle Rabbi, Sajjad Rahmani Dabbagh, Pelin Angin, Ali Kemal Yetisen, and Savas Tasoglu. Deep learning-enabled technologies for bioimage analysis. *Micromachines*, 13(2):260, February 2022.

[6] Travers Ching, Daniel S. Himmelstein, Brett K. Beaulieu-Jones, Alexandr A. Kalinin, Brian T. Do, Gregory P. Way, Enrico Ferrero, Paul-Michael Agapow, Michael Zietz, Michael M. Hoffman, Wei Xie, Gail L. Rosen, Benjamin J. Lengerich, Johnny Israeli, Jack Lanchantin, Stephen Woloszynek, Anne E. Carpenter, Avanti Shrikumar, Jinbo Xu, Evan M. Cofer, Christopher A. Lavender, Srinivas C. Turaga, Amr M. Alexandari, Zhiyong Lu, David J. Harris, Dave DeCaprio, Yanjun Qi, Anshul Kundaje, Yifan Peng, Laura K. Wiley, Marwin H. S. Segler, Simina M. Boca, S. Joshua Swamidass, Austin Huang, Anthony Gitter, and Casey S. Greene. Opportunities and obstacles for deep learning in biology and medicine. *Journal of The Royal Society Interface*, 15(141):20170387, April 2018.

[7] M. Vera, J. Biswas, A. Senecal, R.H. Singer, and H.Y. Park. Single-cell and single-molecule analysis of gene expression regulation. *Annual review of genetics*, 23(50):267–291, 2016.

[8] Andreas Schertel, Nicolas Snaidero, Hong-Mei Han, Torben Ruhwedel, Michael Laue, Markus Grabenbauer, and Wiebke Möbius. Cryo fib-sem: volume imaging of cellular ultrastructure in native frozen specimens. *Journal of structural biology*, 184(2):355—360, November 2013.

[9] K.M Dean and A.E. Palmer. Advances in fluorescence labeling strategies for dynamic cellular imaging. *Nature Chemical Biology*, 10(7):512–23, 2014.

[10] K. Bacia, S.A. Kim, and P. Schwille. Fluorescence cross-correlation spectroscopy in living cells. *Nature Methods*, 3:83–9, 2006.

[11] J. Lippincott-Schwartz and S. Manley. Putting super-resolution fluorescence microscopy to work. *Nature Methods*, 6:21–23, 2009.

[12] V. Micó, J. Zheng, J. Garcia, Z. Zalevsky, and P. Gao. Resolution enhancement in quantitative phase microscopy. *Advances in Optics and Photonics*, 11:135–214, 2019.

[13] H. J. J. BRADDICK. An instrument develops. *Nature*, 215(5106):1205–1206, September 1967.

[14] Michael Davidson and Mortimer Abramowitz. *Optical Microscopy*. 01 2002.

[15] Moritz von Rohr. Theorie und geschichte des photographischen objectivs. page 248, 1899.

[16] Philipp Ludwig von Seidel. *Astr. Nach*, (43):289–304, 1856.

[17] Erin E Diel, Jeff W Lichtman, and Douglas S Richardson. Tutorial: avoiding and correcting sample-induced spherical aberration artifacts in 3d fluorescence microscopy. *Nature protocols*, 15(9):2773–2784, 2020.

[18] José Sasián. *Introduction to aberrations in optical imaging systems*. Cambridge University Press, 2013.

[19] Max Born and Emil Wolf. *Principles of Optics - Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. CUP Archive, Cork, 2000.

[20] von F. Zernike. Beugungstheorie des schneidenver-fahrens und seiner verbesserten form, der phasenkontrastmethode. *Physica*, 1(7-12):689–704, May 1934.

[21] Vasudevan Lakshminarayanan and Andre Fleck. Zernike polynomials: a guide. *Journal of Modern Optics*, 58(18):1678–1678, October 2011.

[22] David S.C. Biggs. A practical guide to deconvolution of fluorescence microscope imagery. *Microscopy Today*, 18(1):10–14, January 2010.

[23] Rudi Rottenfusser. Proper alignment of the microscope. In *Methods in Cell Biology*, pages 43–67. Elsevier, 2013.

[24] Bo Zhang, Josiane Zerubia, and Jean-Christophe Olivo-Marin. Gaussian approximations of fluorescence microscope point-spread function models. *Applied Optics*, 46(10):1819, March 2007.

[25] David J. Stephens and Victoria J. Allan. Light microscopy techniques for live cell imaging. *Science*, 300(5616):82–86, April 2003.

[26] B. Herman. *Fluorescence microscopy*. Springer in Association with the Royal Microscopical Society. Bios Scientific Publishers, 1998.

[27] Marius A. Albota, Chris Xu, and Watt W. Webb. Two-photon fluorescence excitation cross sections of biomolecular probes from 690 to 960 nm. *Applied Optics*, 37(31):7352, November 1998.

[28] A. Jablonski. Efficiency of anti-stokes fluorescence in dyes. *Nature*, 131(3319):839–840, June 1933.

[29] Jeff W Lichtman and José-Angel Conchello. Fluorescence microscopy. *Nature Methods*, 2(12):910–919, November 2005.

[30] Sebastian Dunst and Pavel Tomancak. Imaging flies by fluorescence microscopy: Principles, technologies, and applications. *Genetics*, 211(1):15–34, November 2018.

[31] Christopher P. Toseland. Fluorescent labeling and modification of proteins. *Journal of Chemical Biology*, 6(3):85–95, April 2013.

[32] Jonas Ries, Salvatore Chiantia, and Petra Schwille. Accurate determination of membrane dynamics with line-scan FCS. *Biophysical Journal*, 96(5):1999–2008, March 2009.

[33] Zdeněk Petrášek, Jonas Ries, and Petra Schwille. Scanning FCS for the characterization of protein dynamics in live cells. In *Methods in Enzymology*, pages 317–343. Elsevier, 2010.

[34] Shane M. Hickey, Ben Ung, Christie Bader, Robert Brooks, Joanna Lazniewska, Ian R. D. Johnson, Alexandra Sorvina, Jessica Logan, Carmela Martini, Courtney R. Moore, Litsa Karageorgos, Martin J. Sweetman, and Douglas A. Brooks. Fluorescence microscopy—an outline of hardware, biological handling, and fluorophore considerations. *Cells*, 11(1):35, December 2021.

[35] Kevin L. Holmes and Larry M. Lantz. Chapter 9 protein labeling with fluorescent probes. In *Cytometry*, pages 185–204. Elsevier, 2001.

[36] Yu-Lei Chen, Xin-Xin Xie, Ning Zhong, Le-Chang Sun, Duanquan Lin, Ling-Jing Zhang, Ling Weng, Tengchuan Jin, and Min-Jie Cao. Research progresses and applications of fluorescent protein antibodies: A review focusing on nanobodies. *International Journal of Molecular Sciences*, 24(5):4307, February 2023.

[37] Byron Ballou, B. Christoffer Lagerholm, Lauren A. Ernst, Marcel P. Bruchez, and Alan S. Waggoner. Noninvasive imaging of quantum dots in mice. *Bioconjugate Chemistry*, 15(1):79–86, December 2003.

[38] Quanyu Zhou, Zhenyue Chen, Yu-Hang Liu, Mohamad El Amki, Chaim Glück, Jeanne Droux, Michael Reiss, Bruno Weber, Susanne Wegener, and Daniel Razansky. Three-dimensional wide-field fluorescence microscopy for transcranial mapping of cortical microcirculation. *Nature Communications*, 13(1), December 2022.

[39] Jaroslav Icha, Michael Weber, Jennifer C. Waters, and Caren Norden. Phototoxicity in live fluorescence microscopy, and how to avoid it. *BioEssays*, 39(8):1700003, July 2017.

[40] James B. Pawley, editor. *Handbook Of Biological Confocal Microscopy*. Springer US, 2006.

[41] José-Angel Conchello and Jeff W Lichtman. Optical sectioning microscopy. *Nature Methods*, 2(12):920–931, November 2005.

[42] Kurt Thorn. Spinning-disk confocal microscopy of yeast. In *Methods in Enzymology*, pages 581–602. Elsevier, 2010.

[43] P. Michael Conn. *Imaging and Spectroscopic Analysis of Living Cells - Live Cell Imaging of Cellular Elements and Functions*. Academic Press, Amsterdam, Boston, 2012.

[44] Alexandru Mihai Grumezescu. *Fullerens, Graphenes and Nanotubes - A Pharmaceutical Approach*. William Andrew, Zagreb, 2018.

[45] I Meglinski. *Biophotonics for Medical Applications* -. Elsevier, Amsterdam, 2015.

[46] Michael Schnoor, Lei-Miao Yin, and Sean Sun. *Cell Movement in Health and Disease* -. Academic Press, Amsterdam, Boston, 2022.

[47] Christopher Schmied, Evangelia Stamataki, and Pavel Tomancak. Chapter 27 - open-source solutions for spimage processing. In Jennifer C. Waters and Torsten Wittman, editors, *Quantitative Imaging in Cell Biology*, volume 123 of *Methods in Cell Biology*, pages 505–529. Academic Press, 2014.

[48] Kirti Prakash, Benedict Diederich, Rainer Heintzmann, and Lothar Schermelleh. Super-resolution microscopy: a brief history and new avenues. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 380(2220), February 2022.

[49] Paul Tillberg. Chapter 1 - protein-retention expansion microscopy: Improved subcellular imaging resolution through physical specimen expansion. In Paul Guichard and Virginie Hamel, editors, *Expansion Microscopy for Cell Biology*, volume 161 of *Methods in Cell Biology*, pages 1–14. Academic Press, 2021.

[50] Leonhard Möckl, Don C. Lamb, and Christoph Bräuchle. Super-resolved fluorescence microscopy: Nobel prize in chemistry 2014 for eric betzig, stefan hell, and william e. moerner. *Angewandte Chemie International Edition*, 53(51):13972–13977, November 2014.

[51] M. G. L. Gustafsson. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. SHORT COMMUNICATION. *Journal of Microscopy*, 198(2):82–87, May 2000.

[52] Rainer Heintzmann and Thomas Huser. Super-resolution structured illumination microscopy. *Chemical Reviews*, 117(23):13890–13908, November 2017.

[53] Stefan W. Hell and Jan Wichmann. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics Letters*, 19(11):780, June 1994.

[54] Giuseppe Vicidomini, Paolo Bianchini, and Alberto Diaspro. STED super-resolved microscopy. *Nature Methods*, 15(3):173–182, January 2018.

[55] Eric Betzig, George H. Patterson, Rachid Sougrat, O. Wolf Lindwasser, Scott Olenych, Juan S. Bonifacino, Michael W. Davidson, Jennifer Lippincott-Schwartz, and Harald F. Hess. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642–1645, September 2006.

[56] Samuel T. Hess, Thanu P.K. Girirajan, and Michael D. Mason. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical Journal*, 91(11):4258–4272, December 2006.

[57] Michael J Rust, Mark Bates, and Xiaowei Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods*, 3(10):793–796, August 2006.

[58] Mike Heilemann, Sebastian van de Linde, Mark Schüttpelz, Robert Kasper, Britta Seefeldt, Anindita Mukherjee, Philip Tinnefeld, and Markus Sauer. Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angewandte Chemie International Edition*, 47(33):6172–6176, August 2008.

[59] Mickaël Lelek, Melina T. Gyparaki, Gerti Beliu, Florian Schueder, Juliette Griffié, Suliana Manley, Ralf Jungmann, Markus Sauer, Melike Lakadamyali, and Christophe Zimmer. Single-molecule localization microscopy. *Nature Reviews Methods Primers*, 1(1), June 2021.

[60] Simon J. D. Prince. *Computer Vision - Models, Learning, and Inference*. Cambridge University Press, Cambridge, 2012.

[61] David Forsyth and Jean Ponce. *Computer Vision - A Modern Approach*. Pearson, München, 2012.

[62] Wilhelm Burger and Mark J. Burge. *Principles of Digital Image Processing - Advanced Methods*. Springer Science Business Media, Berlin Heidelberg, 2013.

[63] Rafael C. Gonzalez and Richard Eugene Woods. *Digital Image Processing -*. Prentice Hall, London, 2008.

[64] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, May 2015.

[65] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, May 2017.

[66] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, and Clara I. Sánchez. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42:60–88, December 2017.

[67] George L. Wied, Peter H. Bartels, Marluce Bibbo, and Harvey E. Dytch. Image analysis in quantitative cytopathology and histopathology. *Human Pathology*, 20(6):549–571, June 1989.

[68] B. LERNER, H. GUTERMAN, I. DINSTEIN, and Y. ROMEM. HUMAN CHROMOSOME CLASSIFICATION USING MULTILAYER PERCEPTRON NEURAL NETWORK. *International Journal of Neural Systems*, 06(03):359–370, September 1995.

[69] Per Jesper Sjstrm, Beata Ras Frydel, and Lars Ulrik Wahlberg. Artificial neural network-aided image analysis system for cell counting. *Cytometry*, 36(1):18–26, May 1999.

[70] Kai Huang and Robert F. Murphy. From quantitative microscopy to automated image understanding. *Journal of Biomedical Optics*, 9(5):893, 2004.

[71] Fuyong Xing, Yuanpu Xie, Hai Su, Fujun Liu, and Lin Yang. Deep learning in microscopy image analysis: A survey. *IEEE Transactions on Neural Networks and Learning Systems*, 29(10):4550–4568, October 2018.

[72] Lucas von Chamier, Romain F. Laine, and Ricardo Henriques. Artificial intelligence for microscopy: what you should know. *Biochemical Society Transactions*, 47(4):1029–1040, July 2019.

[73] Jing Sun, Attila Tárnok, and Xuantao Su. Deep learning-based single-cell optical image studies. *Cytometry Part A*, 97(3):226–240, January 2020.

[74] Erik Meijering. A bird's-eye view of deep learning in bioimage analysis. *Computational and Structural Biotechnology Journal*, 18:2312–2325, 2020.

[75] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of machine learning*. Adaptive Computation and Machine Learning series. MIT Press, London, England, 2 edition, December 2018.

[76] Ricard Argelaguet, Anna S. E. Cuomo, Oliver Stegle, and John C. Marioni. Computational principles and challenges in single-cell data integration. *Nature Biotechnology*, 39(10):1202–1215, May 2021.

[77] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.

[78] Thorsten Falk, Dominic Mai, Robert Bensch, Özgün Çiçek, Ahmed Abdulkadir, Yassine Marrakchi, Anton Böhm, Jan Deubner, Zoe Jäckel, Katharina Seiwald, Alexander Dovzhenko, Olaf Tietz, Cristina Dal Bosco, Sean Walsh, Deniz Saltukoglu, Tuan Leng Tay, Marco Prinz, Klaus Palme, Matias Simons, Ilka Diester, Thomas Brox, and Olaf Ronneberger. U-net: deep learning for cell counting, detection, and morphometry. *Nature Methods*, 16(1):67–70, December 2018.

[79] Melissa Linkert, Curtis T. Rueden, Chris Allan, Jean-Marie Burel, Will Moore, Andrew Patterson, Brian Loranger, Josh Moore, Carlos Neves, Donald MacDonald, Aleksandra Tarkowska, Caitlin Sticco, Emma Hill, Mike Rossner, Kevin W. Eliceiri, and Jason R. Swedlow. Metadata matters: access to image data in the real world. *Journal of Cell Biology*, 189(5):777–782, May 2010.

[80] Sébastien Besson, Roger Leigh, Melissa Linkert, Chris Allan, Jean-Marie Burel, Mark Carroll, David Gault, Riad Gozim, Simon Li, Dominik Lindner, Josh Moore, Will Moore, Petr Walczysko, Frances Wong, and Jason R. Swedlow. Bringing open data to whole slide imaging. In *Digital Pathology*, pages 3–10. Springer International Publishing, 2019.

[81] Kota Miura and Nataša Sladoje, editors. *Bioimage Data Analysis Workflows*. Springer International Publishing, 2020.

[82] Andreas Walter, Perrine Paul-Gilloteaux, Birgit Plochberger, Ludek Sefc, Paul Verkade, Julia G. Mannheim, Paul Slezak, Angelika Unterhuber, Martina Marchetti-Deschmann, Manfred Ogris, Katja Bühler, Dror Fixler, Stefan H. Geyer, Wolfgang J. Weninger, Martin Glösmann, Stephan Handschuh, and Thomas Wanek. Correlated multimodal imaging in life sciences: Expanding the biomedical horizon. *Frontiers in Physics*, 8, April 2020.

[83] J Larson. Two-dimensional and three-dimensional blind deconvolution of fluorescence confocal images. In *Proceedings of SPIE*, 2002.

[84] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38:295–307, 2014.

[85] Allan Abraham. Image resolution enhancement system using deep convolutional neural network for image vision applications. *International Journal of Information Systems and Computer Sciences*, 2019.

[86] Min Guo, Yue Li, Yijun Su, Talley Lambert, Damian Dalle Nogare, Mark W. Moyle, Leighton H. Duncan, Richard Ikegami, Anthony Santella, Ivan Rey-Suarez, Daniel Green, Anastasia Beiriger, Jiji Chen, Harshad Vishwasrao, Sundar Ganesan, Victoria Prince, Jennifer C. Waters, Christina M. Annunziata, Markus Hafner, William A. Mohler, Ajay B. Chitnis, Arpita Upadhyaya, Ted B. Usdin, Zhirong Bao, Daniel Colón-Ramos, Patrick La Riviere, Huafeng Liu, Yicong Wu, and Hari Shroff. Rapid image deconvolution and multiview fusion for optical microscopy. *Nature Biotechnology*, 38(11):1337–1346, June 2020.

[87] Abhijeet Rajendra Phatak and Solgaard Lab. 3 d deconvolution with deep learning. 2018.

[88] Yasuhiko Tachibana, Takayuki Obata, Jeff Kershaw, Hironao Sakaki, Takuya Urushihata, Tokuhiko Omatsu, Riwa Kishimoto, and Tatsuya Higashi. The utility of applying various image preprocessing strategies to reduce the ambiguity in deep learning-based clinical image diagnosis. *Magnetic Resonance in Medical Sciences*, 19(2):92–98, 2020.

[89] David Murcia-Gómez, Ignacio Rojas-Valenzuela, and Olga Valenzuela. Impact of image preprocessing methods and deep learning models for classifying histopathological breast cancer images. *Applied Sciences*, 12(22):11375, November 2022.

[90] Massimo Salvi, U. Rajendra Acharya, Filippo Molinari, and Kristen M. Meiburger. The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis. *Computers in Biology and Medicine*, 128:104129, January 2021.

[91] Tim-Oliver Buchholz, Mangal Prakash, Alexander Krull, and Florian Jug. Denoiseg: Joint denoising and segmentation. 2020.

[92] Alexander Krull, Tim-Oliver Buchholz, and Florian Jug. Noise2void-learning denoising from single noisy images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2129–2137, 2019.

[93] Romain F. Laine, Guillaume Jacquemet, and Alexander Krull. Imaging in focus: An introduction to denoising bioimages in the era of deep learning. *The International Journal of Biochemistry &amp Cell Biology*, 140:106077, November 2021.

[94] Linwei Fan, Fan Zhang, Hui Fan, and Caiming Zhang. Brief review of image denoising techniques. *Visual Computing for Industry, Biomedicine, and Art*, 2(1), July 2019.

[95] Ricardo Dutra da Silva, Rodrigo Minetto, William Robson Schwartz, and Helio Pedrini. Adaptive edge-preserving image denoising using wavelet transforms. *Pattern Analysis and Applications*, 16(4):567–580, February 2012.

[96] C. R. Vogel and M. E. Oman. Iterative methods for total variation denoising. *SIAM Journal on Scientific Computing*, 17(1):227–238, January 1996.

[97] Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, December 2010.

[98] Erick Moen, Dylan Bannon, Takamasa Kudo, William Graf, Markus Covert, and David Van Valen. Deep learning for cellular image analysis. *Nature Methods*, 16(12):1233–1246, May 2019.

[99] Martin Weigert, Uwe Schmidt, Tobias Boothe, Andreas Müller, Alexandr Dibrov, Akanksha Jain, Benjamin Wilhelm, Deborah Schmidt, Coleman Broaddus, Siân Culley, Mauricio Rocha-Martins, Fabián Segovia-Miranda, Caren Norden, Ricardo Henriques, Marino Zerial, Michele Solimena, Jochen Rink, Pavel Tomancak, Loic Royer, Florian Jug, and Eugene W. Myers. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nature Methods*, 15(12):1090–1097, November 2018.

[100] Artur Speiser, Lucas-Raphael Müller, Philipp Hoess, Ulf Matti, Christopher J. Obara, Wesley R. Legant, Anna Kreshuk, Jakob H. Macke, Jonas Ries, and Srinivas C. Turaga. Deep learning enables fast and dense single-molecule localization with high accuracy. *Nature Methods*, 18(9):1082–1090, September 2021.

[101] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2noise: Learning image restoration without clean data. 03 2018.

[102] Joshua Batson and Loic Royer. Noise2self: Blind denoising by self-supervision, 2019.

[103] Zhiyong Xie and Gerald Farin. Image registration using hierarchical b-splines. *IEEE transactions on visualization and computer graphics*, 10:85–94, 02 2004.

[104] George Wolberg. *Digital Image Warping*. Systems. IEEE Computer Society Press, Los Alamitos, CA, July 1990.

[105] Suicheng Gu, Xin Meng, Frank C. Sciurba, Hongxia Ma, Joseph Leader, Naftali Kaminski, David Gur, and Jiantao Pu. Bidirectional elastic image registration using b-spline affine transformation. *Computerized Medical Imaging and Graphics*, 38(4):306–314, June 2014.

[106] Fred L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on pattern analysis and machine intelligence*, 11(6):567–585, 1989.

[107] C.O.S. Sorzano, P. Thevenaz, and M. Unser. Elastic registration of biological images using vector-spline regularization. *IEEE Transactions on Biomedical Engineering*, 52(4):652–663, 2005.

[108] C. O. S. Sorzano, M. S. Blagov, P. Thévenaz, E. M. Myasnikova, M. G. Samsonova, and M. Unser. Algorithm for spline-based elastic registration in application to confocal images of gene expression. *Pattern Recognition and Image Analysis*, 16(1):93–96, January 2006.

[109] Z. Wu, Trng Lan, J. Wang, Yang Ding, and Z. Qin. Medical image registration using b-spline transform. *International Journal of Simulation: Systems, Science Technology*, 17:1.1–1.6, 01 2016.

[110] G.E. Christensen and J. He. Consistent nonlinear elastic image registration. In *Proceedings IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA 2001)*, pages 37–43, 2001.

[111] Ignacio Arganda-Carreras, Rodrigo Fernández-González, Arrate Muñoz-Barrutia, and Carlos Ortiz-De-Solorzano. 3d reconstruction of histological sections: Application to mammary gland tissue. *Microscopy Research and Technique*, 73(11):1019–1029, March 2010.

[112] Jennifer C. Waters. Accuracy and precision in quantitative fluorescence microscopy. *Journal of Cell Biology*, 185(7):1135–1148, June 2009.

[113] Erik Meijering. Featurej: A java package for image feature extraction. *ImageJ and Fiji*, 2003.

[114] D. Sage, F.R. Neumann, F. Hediger, S.M. Gasser, and M. Unser. Automatic tracking of individual fluorescence particles: application to the study of chromosome dynamics. *IEEE Transactions on Image Processing*, 14(9):1372–1383, September 2005.

[115] Jean-Yves Tinevez, Nick Perry, Johannes Schindelin, Genevieve M Hoopes, Gregory D Reynolds, Emmanuel Laplantine, Sebastian Y Bednarek, Spencer L Shorte, and Kevin W Eliceiri. Trackmate: An open and extensible platform for single-particle tracking. *Methods*, 115:80–90, 2017.

[116] Claudio Collinet, Martin Stöter, Charles R. Bradshaw, Nikolay Samusik, Jochen C. Rink, Denise Kenski, Bianca Habermann, Frank Buchholz, Robert Henschel, Matthias S. Mueller, Wolfgang E. Nagel, Eugenio Fava, Yannis Kalaidzidis, and Marino Zerial. Systems survey of endocytosis by multiparametric image analysis. *Nature*, 464(7286):243–249, February 2010.

[117] Fernando Amat, William Lemon, Daniel P Mossing, Katie McDole, Yinan Wan, Kristin Branson, Eugene W Myers, and Philipp J Keller. Fast, accurate reconstruction of cell lineages from large-scale fluorescence microscopy data. *Nature Methods*, 11(9):951–958, July 2014.

[118] Seiichi Uchida. Image processing and recognition for biological images. *Development, Growth &amp Differentiation*, 55(4):523–549, April 2013.

[119] Rintu Maria Thomas and Jisha John. A review on cell detection and segmentation in microscopic images. In *2017 International Conference on Circuit ,Power and Computing Technologies (ICCPCT)*, pages 1–5, 2017.

[120] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.

[121] Weiyang Chen, Weiwei Li, Xiangjun Dong, and Jialun Pei. A review of biological image analysis. *Current Bioinformatics*, 13(4):337–343, July 2018.

[122] Ying Yu, Chunping Wang, Qiang Fu, Renke Kou, Fuyu Huang, Boxiong Yang, Tingting Yang, and Mingliang Gao. Techniques and challenges of image segmentation: A review. *Electronics*, 12(5):1199, March 2023.

[123] Michael Held, Michael H A Schmitz, Bernd Fischer, Thomas Walter, Beate Neumann, Michael H Olma, Matthias Peter, Jan Ellenberg, and Daniel W Gerlich. CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging. *Nature Methods*, 7(9):747–754, August 2010.

[124] Jean Serra. Introduction to mathematical morphology. *Computer Vision, Graphics, and Image Processing*, 35(3):283–305, September 1986.

[125] Pierre Soille. *Morphological Image Analysis*. Springer Berlin Heidelberg, 1999.

[126] David Legland, Ignacio Arganda-Carreras, and Philippe Andrey. MorphoLibJ: integrated library and plugins for mathematical morphology with ImageJ. *Bioinformatics*, 32(22):3532–3534, July 2016.

[127] Luis Pedro Coelho, Aabid Shariff, and Robert F. Murphy. Nuclear segmentation in microscope cell images: A hand-segmented dataset and comparison of algorithms. In *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE, June 2009.

[128] R. Delgado-Gonzalo and M. Unser. Spline-based framework for interactive segmentation in biomedical imaging. *IRBM—Ingénierie et Recherche Biomédicale / BioMedical Engineering and Research*, 34(3):235–243, June 2013.

[129] A Dufour, R Thibeaux, E Labruyere, N Guillen, and J-C Olivo-Marin. 3-d active meshes: Fast discrete deformable models for cell tracking in 3-d time-lapse microscopy. *IEEE Transactions on Image Processing*, 20(7):1925–1937, July 2011.

[130] Ricard Delgado-Gonzalo, Nicolas Chenouard, and Michael Unser. Fast parametric snakes for 3d microscopy. In *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*. IEEE, May 2012.

[131] Ricard Delgado-Gonzalo, Nicolas Chenouard, and Michael Unser. Spline-based deforming ellipsoids for interactive 3d bioimage segmentation. *IEEE Transactions on Image Processing*, 22(10):3926–3940, October 2013.

[132] Ian H Witten, Eibe Frank, Mark A Hall, and Christopher J Pal. What's it all about? In *Data Mining*, pages 3–41. Elsevier, 2017.

[133] Ignacio Arganda-Carreras, Verena Kaynig, Curtis Rueden, Kevin W Eliceiri, Johannes Schindelin, Albert Cardona, and H Sebastian Seung. Trainable weka segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics*, 33(15):2424–2426, March 2017.

[134] A. K. Jain, M. N. Murty, and P. J. Flynn. Data clustering. *ACM Computing Surveys*, 31(3):264–323, September 1999.

[135] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2980–2988, 2017.

[136] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection, 2015.

[137] Zhong-Qiu Zhao, Peng Zheng, Shou-Tao Xu, and Xindong Wu. Object detection with deep learning: A review. *IEEE Transactions on Neural Networks and Learning Systems*, 30(11):3212–3232, 2019.

[138] Uwe Schmidt, Martin Weigert, Coleman Broaddus, and Gene Myers. Cell detection with star-convex polygons. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, pages 265–273. Springer International Publishing, 2018.

[139] Soham Mandal and Virginie Uhlmann. Splinedist: Automated cell segmentation with spline curves. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pages 1082–1086, 2021.

[140] Juan C. Caicedo, Allen Goodman, Kyle W. Karhohs, Beth A. Cimini, Jeanelle Ackerman, Marzieh Haghighi, CherKeng Heng, Tim Becker, Minh Doan, Claire McQuin, Mohammad Rohban, Shantanu Singh, and Anne E. Carpenter. Nucleus segmentation across imaging experiments: the 2018 data science bowl. *Nature Methods*, 16(12):1247–1253, October 2019.

[141] Carsen Stringer, Tim Wang, Michalis Michaelos, and Marius Pachitariu. Cellpose: a generalist algorithm for cellular segmentation. *Nature Methods*, 18(1):100–106, December 2020.

[142] Giona Kleinberg, Sophia Wang, Ester Comellas, James R Monaghan, and Sandra J Shefelbine. Usability of deep learning pipelines for 3d nuclei identification with stardist and cellpose. *Cells & Development*, 172:203806, 2022.

[143] Ariel Waisman, Alessandra Marie Norris, Martín Elías Costa, and Daniel Kopinke. Automatic and unbiased segmentation and quantification of myofibers in skeletal muscle. *Scientific reports*, 11(1):11793, 2021.

[144] Nicola Bougen-Zhukov, Sheng Yang Loh, Hwee Kuan Lee, and Lit-Hsin Loo. Large-scale image-based screening and profiling of cellular phenotypes. *Cytometry Part A*, 91(2):115–125, July 2016.

[145] Karsten Rodenacker and Ewert Bengtsson. A feature set for cytometry on digitized microscopic images. *Analytical Cellular Pathology*, 25(1):1–36, 2003.

[146] Juan C Caicedo, Sam Cooper, Florian Heigwer, Scott Warchal, Peng Qiu, Csaba Molnar, Aliaksei S Vasilevich, Joseph D Barry, Harmanjit Singh Bansal, Oren Kraus, Mathias Wawer, Lassi Paavolainen, Markus D Herrmann, Mohammad Rohban, Jane Hung, Holger Hennig, John Concannon, Ian Smith, Paul A Clemons, Shantanu Singh, Paul Rees, Peter Horvath, Roger G Linington, and Anne E Carpenter. Data-analysis strategies for image-based cell profiling. *Nature Methods*, 14(9):849–863, September 2017.

[147] Nilesh Bhaskarrao Bahadure, Arun Kumar Ray, and Har Pal Thethi. Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM. *International Journal of Biomedical Imaging*, 2017:1–12, 2017.

[148] Tommy Löfstedt, Patrik Brynolfsson, Thomas Asklund, Tufve Nyholm, and Anders Garpebring. Gray-level invariant haralick texture features. *PLOS ONE*, 14(2):e0212110, February 2019.

[149] Robert M Haralick, Karthikeyan Shanmugam, and Its' Hak Dinstein. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6):610–621, 1973.

[150] Subrajeet Mohapatra and Dipti Patra. Automated cell nucleus segmentation and acute leukemia detection in blood microscopic images. In *2010 International Conference on Systems in Medicine and Biology*, pages 49–54, 2010.

[151] Carlos Affonso, André Luis Debiaso Rossi, Fábio Henrique Antunes Vieira, and André Carlos Ponce de Leon Ferreira de Carvalho. Deep learning for biological image classification. *Expert Systems with Applications*, 85:114–122, November 2017.

[152] Ismail Khalid Kazmi, Lihua You, and Jian Jun Zhang. A survey of 2d and 3d shape descriptors. In *2013 10th International Conference Computer Graphics, Imaging and Visualization*, pages 1–10. IEEE, 2013.

[153] Huw D. Summers, John W. Wills, and Paul Rees. Spatial statistics is a comprehensive tool for quantifying cell neighbor relationships and biological processes via tissue image analysis. *Cell Reports Methods*, 2(11):100348, November 2022.

[154] Christoph Sommer and Daniel W. Gerlich. Machine learning in cell biology – teaching computers to recognize phenotypes. *Journal of Cell Science*, January 2013.

[155] Pegah Khosravi, Ehsan Kazemi, Qiansheng Zhan, Jonas E. Malmsten, Marco Toschi, Pantelis Zisimopoulos, Alexandros Sigaras, Stuart Lavery, Lee A. D. Cooper, Cristina Hickman, Marcos Meseguer, Zev Rosenwaks, Olivier Elemento, Nikica Zaninovic, and Iman Hajirasouliha. Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization. *npj Digital Medicine*, 2(1), April 2019.

[156] Samuel J. Yang, Marc Berndl, D. Michael Ando, Mariya Barch, Arunachalam Narayanaswamy, Eric Christiansen, Stephan Hoyer, Chris Roat, Jane Hung, Curtis T. Rueden, Asim Shankar, Steven Finkbeiner, and Philip Nelson. Assessing microscope image focus quality with deep learning. *BMC Bioinformatics*, 19(1), March 2018.

[157] Fengyu Cong, Sijie Lin, Hongkai Wang, Shang Shang, Ling Long, Ruxue Hu, Yi Wu, Na Chen, and Shaoxiang Zhang. Biological image analysis using deep learning-based methods: Literature review. *Digital Medicine*, 4(4):157, 2018.

[158] Oren Z. Kraus, Jimmy Lei Ba, and Brendan J. Frey. Classifying and segmenting microscopy images with deep multiple instance learning. *Bioinformatics*, 32(12):i52–i59, June 2016.

[159] Devin P Sullivan, Casper F Winsnes, Lovisa Åkesson, Martin Hjelmare, Mikaela Wiking, Rutger Schutten, Linzi Campbell, Hjalti Leifsson, Scott Rhodes, Andie Nordgren, Kevin Smith, Bernard Revaz, Bergur Finnbogason, Attila Szantner, and Emma Lundberg. Deep learning is combined with massive-scale citizen science to improve large-scale image classification. *Nature Biotechnology*, 36(9):820–828, October 2018.

[160] Mohammad Shifat-E-Rabbi, Xuwang Yin, Cailey E. Fitzgerald, and Gustavo K. Rohde. Cell image classification: A comparative overview. *Cytometry Part A*, 97(4):347–362, February 2020.

[161] Mojca Mattiazzi Usaj, Erin B. Styles, Adrian J. Verster, Helena Friesen, Charles Boone, and Brenda J. Andrews. High-content screening for quantitative cell biology. *Trends in Cell Biology*, 26(8):598–611, August 2016.

[162] Ben T. Grys, Dara S. Lo, Nil Sahin, Oren Z. Kraus, Quaid Morris, Charles Boone, and Brenda J. Andrews. Machine learning and computer vision approaches for phenotypic profiling. *Journal of Cell Biology*, 216(1):65–71, December 2016.

[163] Juan C Caicedo, Shantanu Singh, and Anne E Carpenter. Applications in image-based profiling of perturbations. *Current Opinion in Biotechnology*, 39:134–142, June 2016.

[164] Carla Tameling, Stefan Stoldt, Till Stephan, Julia Naas, Stefan Jakobs, and Axel Munk. Colocalization for super-resolution microscopy via optimal transport. *Nature Computational Science*, 1(3):199–211, March 2021.

[165] Kenneth W. Dunn, Malgorzata M. Kamocka, and John H. McDonald. A practical guide to evaluating colocalization in biological microscopy. *American Journal of Physiology-Cell Physiology*, 300(4):C723–C742, April 2011.

[166] M M Mukaka. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med. J.*, 24(3):69–71, September 2012.

[167] Jeremy Adler and Ingela Parmryd. Quantifying colocalization: The case for discarding the scpmanders/scp overlap coefficient. *Cytometry Part A*, 99(9):910–920, March 2021.

[168] Jeremy Adler and Ingela Parmryd. Quantifying colocalization by correlation: The pearson correlation coefficient is superior to the mander's overlap coefficient. *Cytometry Part A*, 77A(8):733–742, March 2010.

[169] Vadim Zinchuk, Olga Zinchuk, and Teruhiko Okada. Quantitative colocalization analysis of multicolor confocal immunofluorescence microscopy images: Pushing pixels to explore biological phenomena. *ACTA HISTOCHEMICA ET CYTOCHEMICA*, 40(4):101–111, 2007.

[170] Jesse S. Aaron, Aaron B. Taylor, and Teng-Leong Chew. Image co-localization – co-occurrence versus correlation. *Journal of Cell Science*, 131(3), February 2018.

[171] Sebastian Malkusch, Ulrike Endesfelder, Justine Mondry, Márton Gelléri, Peter J. Verveer, and Mike Heilemann. Coordinate-based colocalization analysis of single-molecule localization microscopy data. *Histochemistry and Cell Biology*, 137(1):1–10, November 2011.

[172] B. D. Ripley. The second-order analysis of stationary point processes. *Journal of Applied Probability*, 13(2):255–266, June 1976.

[173] Khuloud Jaqaman, Dinah Loerke, Marcel Mettlen, Hirotaka Kuwata, Sergio Grinstein, Sandra L Schmid, and Gaudenz Danuser. Robust single-particle tracking in live-cell time-lapse sequences. *Nature methods*, 5(8):695–702, 2008.

[174] Lei Yang, Zhen Qiu, Alan H Greenaway, and Weiping Lu. A new framework for particle detection in low-snr fluorescence live-cell images and its application for improved particle tracking. *IEEE transactions on biomedical engineering*, 59(7):2040–2050, 2012.

[175] Pascal Vallotton and Sandra Olivier. Tri-track: free software for large-scale particle tracking. *Microscopy and Microanalysis*, 19(2):451–460, 2013.

[176] Nicolas Chenouard, Isabelle Bloch, and Jean-Christophe Olivo-Marin. Multiple hypothesis tracking for cluttered biological image sequences. *IEEE transactions on pattern analysis and machine intelligence*, 35(11):2736–3750, 2013.

[177] Bo Shuang, Jixin Chen, Lydia Kisley, and Christy F Landes. Troika of single particle tracking programing: Snr enhancement, particle identification, and mapping. *Physical Chemistry Chemical Physics*, 16(2):624–634, 2014.

[178] Liang Liang, Hongying Shen, Pietro De Camilli, and James S Duncan. A novel multiple hypothesis based particle tracking method for clathrin mediated endocytosis analysis using fluorescence microscopy. *IEEE transactions on image processing*, 23(4):1844–1857, 2014.

[179] Nicolas Chenouard, Ihor Smal, Fabrice De Chaumont, Martin Maška, Ivo F Sbalzarini, Yuanhao Gong, Janick Cardinale, Craig Carthel, Stefano Coraluppi, Mark Winter, et al. Objective comparison of particle tracking methods. *Nature methods*, 11(3):281–289, 2014.

[180] Astha Jaiswal, William J Godinez, Roland Eils, Maik Jörg Lehmann, and Karl Rohr. Tracking virus particles in fluorescence microscopy images using multi-scale detection and multi-frame association. *IEEE Transactions on Image Processing*, 24(11):4122–4136, 2015.

[181] Ihor Smal and Erik Meijering. Quantitative comparison of multiframe data association techniques for particle tracking in time-lapse fluorescence microscopy. *Medical image analysis*, 24(1):163–189, 2015.

[182] Vladimír Ulman, Martin Maška, Klas E G Magnusson, Olaf Ronneberger, Carsten Haubold, Nathalie Harder, Pavel Matula, Petr Matula, David Svoboda, Miroslav Radojevic, Ihor Smal, Karl Rohr, Joakim Jaldén, Helen M Blau, Oleh Dzyubachyk, Boudewijn Lelieveldt, Pengdong Xiao, Yuexiang Li, Siu-Yeung Cho, Alexandre C Dufour, Jean-Christophe Olivo-Marin, Constantino C Reyes-Aldasoro, Jose A Solis-Lemus, Robert Bensch, Thomas Brox, Johannes Stegmaier, Ralf Mikut, Steffen Wolf, Fred A Hamprecht, Tiago Esteves, Pedro Quelhas, Ömer Demirel, Lars Malmström, Florian Jug, Pavel Tomancak, Erik Meijering, Arrate Muñoz-Barrutia, Michal Kozubek, and Carlos Ortiz de Solorzano. An objective comparison of cell-tracking algorithms. *Nature Methods*, 14(12):1141–1152, October 2017.

[183] Philipp Eulenberg, Niklas Köhler, Thomas Blasi, Andrew Filby, Anne E. Carpenter, Paul Rees, Fabian J. Theis, and F. Alexander Wolf. Reconstructing cell cycle and disease progression using deep learning. *Nature Communications*, 8(1), September 2017.

[184] Martin Hailstone, Dominic Waithe, Tamsin J Samuels, Lu Yang, Ita Costello, Yoav Arava, Elizabeth Robertson, Richard M Parton, and Ilan Davis. CytoCensus, mapping cell identity and division in tissues and organs using machine learning. *eLife*, 9, May 2020.

[185] Chentao Wen, Takuya Miura, Venkatakaushik Voleti, Kazushi Yamaguchi, Motosuke Tsutsumi, Kei Yamamoto, Kohei Otomo, Yukako Fujie, Takayuki Teramoto, Takeshi Ishihara, Kazuhiro Aoki, Tomomi Nemoto, Elizabeth MC Hillman, and Koutarou D Kimura. 3deecelltracker, a deep learning-based pipeline for segmenting and tracking cells in 3d time lapse images. *eLife*, 10, March 2021.

[186] Alexander Mathis, Pranav Mamidanna, Kevin M. Cury, Taiga Abe, Venkatesh N. Murthy, Mackenzie Weygandt Mathis, and Matthias Bethge. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, 21(9):1281–1289, August 2018.

[187] Yao Yao, Ihor Smal, Ilya Grigoriev, Anna Akhmanova, and Erik Meijering. Deep-learning method for data association in particle tracking. *Bioinformatics*, 36(19):4935–4941, July 2020.

[188] Lucas von Chamier, Romain F. Laine, Johanna Jukkala, Christoph Spahn, Daniel Krentzel, Elias Nehme, Martina Lerche, Sara Hernández-Pérez, Pieta K. Mattila, Eleni Karinou, Séamus Holden, Ahmet Can Solak, Alexander Krull, Tim-Oliver Buchholz, Martin L. Jones, Loïc A. Royer, Christophe Leterrier, Yoav Shechtman, Florian Jug, Mike Heilemann, Guillaume Jacquemet, and Ricardo Henriques. Democratising deep

learning for microscopy with ZeroCostDL4mic. *Nature Communications*, 12(1), April 2021.

[189] Thorsten Wagner, Alexandra Kroll, Chandrashekara Haramagatti, Hans-Gerd Lipinski, and Martin Wiemann. Classification and segmentation of nanoparticle diffusion trajectories in cellular micro environments. *PLoS ONE*, 12:e0170165, 01 2017.

[190] Gang Ruan, Amit Agrawal, Adam I. Marcus, and Shuming Nie. Imaging and tracking of tat peptide-conjugated quantum dots in living cells: new insights into nanoparticle uptake, intracellular transport, and vesicle shedding. *Journal of the American Chemical Society*, 129(47):14759–14766, November 2007.

[191] Nilah Monnier, Syuan-Ming Guo, Masashi Mori, Jun He, Péter Lénárt, and Mark Bathe. Bayesian approach to MSD-based analysis of particle motion in live cells. *Biophysical Journal*, 103(3):616–626, August 2012.

[192] Felix Höfling and Thomas Franosch. Anomalous transport in the crowded world of biological cells. *Reports on Progress in Physics*, 76(4):046602, March 2013.

[193] R. Ferrari, A.J. Manfroi, and W.R. Young. Strongly and weakly self-similar diffusion. *Physica D: Nonlinear Phenomena*, 154(1-2):111–137, 2001.

[194] H. Ewers, A.E. Smith, I.F. Sbalzarini, H. Lilie, P. Koumoutsakos, and A. Helenius. Single-particle tracking of murine polyoma virus-like particles on live cells and artificial membranes. *Proceedings of the National Academy of Sciences*, 102(42), 2005.

[195] J.P. Siebrasse, I. Djuric, U. Schulze, M.A. Schlüter, H. Pavenstädt, T. Weide, and U. Kubitscheck. Trajectories and single-particle tracking data of intracellular vesicles loaded with either snap-crb3a or snap-crb3b. *Data in Brief*, 7(9465):1665–1669, 2016.

[196] Jianqing Fan, Fang Han, and Han Liu. Challenges of big data analysis. *National Science Review*, 1(2):293–314, February 2014.

[197] Jennifer Harrow, John Hancock, Niklas Blomberg, Niklas Blomberg, Søren Brunak, Salvador Capella-Gutierrez, Christine Durinx, Chris T. Evelo, Carole Goble, Ivo Gut, Jon Ison, Thomas Keane, Brane Leskošek, Luděk Matyska, Johanna McEntyre, Célia Miguel, Arcadi Navarro, Steven Newhouse, Tommi Nyrönen, Patricia Palagi, Bengt Persson, Cyril Pommier, Jordi Rambla, Marco Roos, Gabriella Rustici, Andrew Smith, Alfonso Valencia, Celia van Gelder, Jiri Vondrasek, Nils Peder Willassen, Juan Arenas, Helen Parkinson, Robert D Finn, Sergi Beltran, Leslie Matalonga, Hannah Hurst, Paul Kersey, Ilkka Lappalainen, Pascal Kahlem, Gary Saunders, Sirarat Sarntivijai, Rachel Drysdale, Johnathan Tedds, Jeremy Lanfear, and Jennifer Harrow and. ELIXIR-EXCELERATE: establishing europe's data infrastructure for the life science research of the future. *The EMBO Journal*, 40(6), February 2021.

[198] Perrine Paul-Gilloteaux, Sébastien Tosi, Jean-Karim Hériché, Alban Gaignard, Hervé Ménager, Raphaël Marée, Volker Baecker, Anna Klemm, Matúš Kalaš, Chong Zhang, Kota Miura, and Julien Colombelli. Bioimage analysis workflows: community resources to navigate through a complex ecosystem. *F1000Research*, 10:320, April 2021.

[199] Jeremy Leipzig. A review of bioinformatic pipeline frameworks. *Briefings in Bioinformatics*, page bbw020, March 2016.

[200] Vahid Jalili, Enis Afgan, Qiang Gu, Dave Clements, Daniel Blankenberg, Jeremy Goecks, James Taylor, and Anton Nekrutenko. The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. *Nucleic Acids Research*, 48(W1):W395–W402, June 2020.

[201] Philip A. Ewels, Alexander Peltzer, Sven Fillinger, Harshil Patel, Johannes Alneberg, Andreas Wilm, Maxime Ulysse Garcia, Paolo Di Tommaso, and Sven Nahnsen. The nf-core framework for community-curated bioinformatics pipelines. *Nature Biotechnology*, 38(3):276–278, February 2020.

[202] Ulysse Rubens, Romain Mormont, Lassi Paavolainen, Volker Bäcker, Benjamin Pavie, Leandro A. Scholz, Gino Michiels, Martin Maška, Devrim Ünay, Graeme Ball, Renaud Hoyoux, Rémy Vandaele, Ofra Golani, Stefan G. Stanciu, Natasa Sladoje, Perrine Paul-Gilloteaux, Raphaël Marée, and Sébastien Tosi. BIAFLOWS: A collaborative framework to reproducibly deploy and benchmark bioimage analysis workflows. *Patterns*, 1(3):100040, June 2020.

[203] Christopher Schmied, Peter Steinbach, Tobias Pietzsch, Stephan Preibisch, and Pavel Tomancak. An automated workflow for parallel processing of large multiview SPIM recordings. *Bioinformatics*, 32(7):1112–1114, December 2015.

[204] Robert Haase, Loic A. Royer, Peter Steinbach, Deborah Schmidt, Alexandr Dibrov, Uwe Schmidt, Martin Weigert, Nicola Maghelli, Pavel Tomancak, Florian Jug, and Eugene W. Myers. CLIJ: GPU-accelerated image processing for everyone. *Nature Methods*, 17(1):5–6, November 2019.

[205] Ibrahim Masri and Erdal Erdal. Review paper on real time image processing: Methods, techniques, applications. 06 2019.

[206] P Dillinger, JF Vogelbruch, J Leinen, S Suslov, R Patzak, H Winkler, and K Schwan. Fpga based real-time image segmentation for medical systems and data processing. In *14th IEEE-NPSS Real Time Conference, 2005.*, pages 5–pp. IEEE, 2005.

[207] Sparsh Mittal, Saket Gupta, and Sudeb Dasgupta. Fpga: An efficient and promising platform for real-time image processing applications. In *National Conference On Research and Development In Hardware Systems (CSI-RDHS)*, 2008.

[208] S Mori. Real-time image-processing algorithm for markerless tumour tracking using x-ray fluoroscopic imaging. *The British Journal of Radiology*, 87(1037):20140001, May 2014.

[209] Peng Meng, Hui Zhang, Haiying Wen, Dai Min, Zhisheng Zhang, and Guifu Zhang. Design of real-time image processing system of medical high-definition electronic endoscope based on fpga. In *2022 28th International Conference on Mechatronics and Machine Vision in Practice (M2VIP)*, pages 1–6, 2022.

[210] M Fathy and MY Siyal. Real-time image processing approach to measure traffic queue parameters. *IEE Proceedings-Vision, Image and Signal Processing*, 142(5):297–303, 1995.

[211] Naeem Abbas, Muhammad Tayyab, and M Tahir Qadri. Real time traffic density count using image processing. *International Journal of Computer Applications*, 83(9):16–19, 2013.

[212] Xavier P Burgos-Artizzu, Angela Ribeiro, Maria Guijarro, and Gonzalo Pajares. Real-time image processing for crop/weed discrimination in maize fields. *Computers and Electronics in Agriculture*, 75(2):337–346, 2011.

[213] S Maheswaran, M Ramya, P Priyadharshini, and P Sivaranjani. A real time image processing based system to scaring the birds from the agricultural field. *Indian Journal of Science and Technology*, 9(30):1–5, 2016.

[214] Jong Jae Lee and Masanobu Shinozuka. A vision-based system for remote sensing of bridge displacement. *Ndt & E International*, 39(5):425–431, 2006.

[215] Baogui Qi, Hao Shi, Yin Zhuang, He Chen, and Liang Chen. On-board, real-time preprocessing system for optical remote-sensing imagery. *Sensors*, 18(5):1328, 2018.

[216] Chen Chen, Wei Li, Lianru Gao, Hengchao Li, and Javier Plaza. Special issue on advances in real-time image processing for remote sensing. *Journal of Real-Time Image Processing*, 15:435–438, 2018.

[217] J. Gómez-Blanco, J.M. de la Rosa-Trevín, R. Marabini, L. del Cano, A. Jiménez, M. Martínez, R. Melero, T. Majtner, D. Maluenda, J. Mota, Y. Rancel, E Ramírez-Aportela, J.L. Vilas, M. Carroni, S. Fleischmann, E. Lindahl, A.W. Ashton, M. Basham, D.K. Clare, K. Savage, C.A. Siebert, G.G. Sharov, C.O.S. Sorzano, P. Conesa, and J.M. Carazo. Using scipion for stream image processing at cryo-EM facilities. *Journal of Structural Biology*, 204(3):457–463, December 2018.

[218] D. Maluenda, T. Majtner, P. Horvath, J. L. Vilas, A. Jiménez-Moreno, J. Mota, E. Ramírez-Aportela, R. Sánchez-García, P. Conesa, L. del Caño, Y. Rancel, Y. Fonseca, M. Martínez, G. Sharov, C.A. García, D. Strelak, R. Melero, R. Marabini, J. M. Carazo, and C. O. S. Sorzano. Flexible workflows for on-the-fly electron-microscopy single-particle image processing using iscipion/i. *Acta Crystallographica Section D Structural Biology*, 75(10):882–894, October 2019.

[219] Young Jin Heo, Donghyeon Lee, Junsu Kang, Keondo Lee, and Wan Kyun Chung. Real-time image processing for microscopy-based label-free imaging flow cytometry in a microfluidic chip. *Scientific Reports*, 7(1), September 2017.

[220] Xinyang Li, Yixin Li, Yiliang Zhou, Jiamin Wu, Zhifeng Zhao, Jiaqi Fan, Fei Deng, Zhaofa Wu, Guihua Xiao, Jing He, Yuanlong Zhang, Guoxun Zhang, Xiaowan Hu, Xingye Chen, Yi Zhang, Hui Qiao, Hao Xie, Yulong Li, Haoqian Wang, Lu Fang, and Qionghai Dai. Real-time denoising enables high-sensitivity fluorescence time-lapse imaging beyond the shot-noise limit. *Nature Biotechnology*, 41(2):282–292, September 2022.

[221] Alex Gaudeul. Do open source developers respond to competition? the LATEX case study. *Review of Network Economics*, 6(2), January 2007.

[222] The gnu operating system. https://www.gnu.org/. Accessed: 20223-05-13.

[223] History of the osi. https://opensource.org/history/. Accessed: 20223-05-13.

[224] Romain Guiet, Olivier Burri, and Arne Seitz. Open source tools for biological image analysis. In *Computer Optimized Microscopy*, pages 23–37. Springer New York, 2019.

[225] Caroline A Schneider, Wayne S Rasband, and Kevin W Eliceiri. NIH image to ImageJ: 25 years of image analysis. *Nature Methods*, 9(7):671–675, June 2012.

[226] Curtis T. Rueden, Johannes Schindelin, Mark C. Hiner, Barry E. DeZonia, Alison E. Walter, Ellen T. Arena, and Kevin W. Eliceiri. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics*, 18(1), November 2017.

[227] Mark C. Hiner, Curtis T. Rueden, and Kevin W. Eliceiri. SCIFIO: an extensible framework to support scientific image formats. *BMC Bioinformatics*, 17(1), December 2016.

[228] Anne E Carpenter, Thouis R Jones, Michael R Lamprecht, Colin Clarke, In Kang, Ola Friman, David A Guertin, Joo Chang, Robert A Lindquist, Jason Moffat, Polina Golland, and David M Sabatini. *Genome Biology*, 7(10):R100, 2006.

[229] David R Stirling, Anne E Carpenter, and Beth A Cimini. CellProfiler analyst 3.0: accessible data exploration and machine learning for image analysis. *Bioinformatics*, 37(21):3992–3994, September 2021.

[230] Marcel Austenfeld and Wolfram Beyschlag. A graphical user interface forir/iin a rich client platform for ecological modeling. *Journal of Statistical Software*, 49(4), 2012.

[231] Rosa Doran, Anne-Laure Melchior, Thomas Boudier, Pacôme, Delva, Roger Ferlet, Maria L. T. Almeida, Domingos Barbosa, Edward Gomez, Carl Pennypacker, Paul Roche, and Sarah Roberts. Astrophysics datamining in the classroom: Exploring real data with new software tools and robotic telescopes, 2012.

[232] Karen A. Collins, John F. Kielkopf, Keivan G. Stassun, and Frederic V. Hessman. ASTROIMAGEJ: IMAGE PROCESSING AND PHOTOMETRIC EXTRACTION FOR ULTRA-PRECISE ASTRONOMICAL LIGHT CURVES. *The Astronomical Journal*, 153(2):77, January 2017.

[233] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021.

[234] Fabrice de Chaumont, Stéphane Dallongeville, Nicolas Chenouard, Nicolas Hervé, Sorin Pop, Thomas Provoost, Vannary Meas-Yedid, Praveen Pankajakshan, Timothée Lecomte, Yoann Le Montagner, Thibault Lagache, Alexandre Dufour, and Jean-Christophe Olivo-Marin. Icy: an open bioimage informatics platform for extended reproducible research. *Nature Methods*, 9(7):690–696, June 2012.

[235] Peter Bankhead, Maurice B. Loughrey, José A. Fernández, Yvonne Dombrowski, Darragh G. McArt, Philip D. Dunne, Stephen McQuaid, Ronan T. Gray, Liam J. Murray, Helen G. Coleman, Jacqueline A. James, Manuel Salto-Tellez, and Peter W. Hamilton. QuPath: Open source software for digital pathology image analysis. *Scientific Reports*, 7(1), December 2017.

[236] Michael R. Berthold, Nicolas Cebron, Fabian Dill, Thomas R. Gabriel, Tobias Kötter, Thorsten Meinl, Peter Ohl, Christoph Sieb, Kilian Thiel, and Bernd Wiswedel. KNIME: The konstanz information miner. In *Data Analysis, Machine Learning and Applications*, pages 319–326. Springer Berlin Heidelberg, 2008.

[237] Stuart Berg, Dominik Kutra, Thorben Kroeger, Christoph N. Straehle, Bernhard X. Kausler, Carsten Haubold, Martin Schiegg, Janez Ales, Thorsten Beier, Markus Rudy, Kemal Eren, Jaime I. Cervantes, Buote Xu, Fynn Beuttenmueller, Adrian Wolny, Chong Zhang, Ullrich Koethe, Fred A. Hamprecht, and Anna Kreshuk. ilastik: interactive machine learning for (bio)image analysis. *Nature Methods*, September 2019.

[238] Estibaliz Gómez de Mariscal, Carlos García-López de Haro, Wei Ouyang, Laurène Donati, Emma Lundberg, Michael Unser, Arrate Muñoz-Barrutia, and Daniel Sage. DeepImageJ: A user-friendly environment to run deep learning models in ImageJ. October 2019.

[239] Wei Ouyang, Fynn Beuttenmueller, Estibaliz Gómez de Mariscal, Constantin Pape, Tom Burke, Carlos Garcia-López de Haro, Craig Russell, Lucía Moya-Sans, Cristina de-la Torre-Gutiérrez, Deborah Schmidt, Dominik Kutra, Maksim Novikov, Martin Weigert, Uwe Schmidt, Peter Bankhead, Guillaume Jacquemet, Daniel Sage, Ricardo Henriques, Arrate Muñoz-Barrutia, Emma Lundberg, Florian Jug, and Anna Kreshuk. BioImage model zoo: A community-driven resource for accessible deep learning in BioImage analysis. June 2022.

[240] Réka Hollandi, Ákos Diósdi, Gábor Hollandi, Nikita Moshkov, and Péter Horváth. AnnotatorJ: an ImageJ plugin to ease hand annotation of cellular compartments. *Molecular Biology of the Cell*, 31(20):2179–2186, September 2020.

[241] Florian Fuchs, Gregoire Pau, Dominique Kranz, Oleg Sklyar, Christoph Budjan, Sandra Steinbrink, Thomas Horn, Angelika Pedal, Wolfgang Huber, and Michael Boutros. Clustering phenotype populations by genome-wide RNAi and multiparametric imaging. *Molecular Systems Biology*, 6(1), January 2010.

[242] Ana Cayuela López, José A. Gómez-Pedrero, Ana M. O. Blanco, and Carlos Oscar S. Sorzano. Cell-TypeAnalyzer: A flexible fiji/ImageJ plugin to classify cells according to user-defined criteria. *Biological Imaging*, 2, 2022.

[243] Laura Martínez-Muñoz, José Miguel Rodríguez-Frade, Rubén Barroso, Carlos Óscar S. Sorzano, Juan A. Torreño-Pina, César A. Santiago, Carlo Manzo, Pilar Lucas, Eva M. García-Cuesta, Enric Gutierrez, Laura Barrio, Javier Vargas, Graciela Cascio, Yolanda R. Carrasco, Francisco Sánchez-Madrid, María F. García-Parajo, and Mario Mellado. Separating actin-dependent chemokine receptor nanoclustering from dimerization indicates a role for clustering in CXCR4 signaling and function. *Molecular Cell*, 70(1):106–119.e10, April 2018.

[244] Ana Cayuela López, Eva M. García-Cuesta, Sofía R. Gardeta, José Miguel Rodríguez-Frade, Mario Mellado, José Antonio Gómez-Pedrero, and Carlos Oscar S. Sorzano. Trackanalyzer: A fiji/imagej toolbox for a holistic analysis of tracks. *Biological Imaging*, page 1–14, 2023.

[245] Ana Cayuela López, Pablo Conesa, Ana María Oña Blanco, José Antonio Gómez-Pedrero, and Carlos Oscar S Sorzano. Real-time correction of chromatic aberration in optical fluorescence microscopy. *Methods and Applications in Fluorescence*, 11(4):045001, July 2023.

Appendix A

Cell-TypeAnalyzer: A flexible Fiji/ImageJ plugin to classify cells according to user-defined criteria

Cayuela López, A., Gómez-Pedrero, J., Blanco, A., Sorzano, C. (2022). Cell-TypeAnalyzer: A flexible Fiji/ImageJ plugin to classify cells according to user-defined criteria. *Biological Imaging*, 2, E5. doi:10.1017/S2633903X22000058

Biological Imaging (2022), 2: e5
doi:10.1017/S2633903X22000058

CAMBRIDGE UNIVERSITY PRESS

SOFTWARE REPORT

Cell-TypeAnalyzer: A flexible Fiji/ImageJ plugin to classify cells according to user-defined criteria

Ana Cayuela López[1]*, José A. Gómez-Pedrero[2], Ana M. O. Blanco[3] and Carlos Oscar S. Sorzano[1]*

[1]Biocomputing Unit, National Centre for Biotechnology, Madrid, Spain
[2]Applied Optics Complutense Group, Faculty of Optics and Optometry, University Complutense of Madrid, Madrid, Spain
[3]Advanced Light Microscopy Unit, National Centre for Biotechnology, Madrid, Spain
*Corresponding authors. E-mail: acayuela@cnb.csic.es; coss@cnb.csic.es

Received: 15 December 2021; **Revised:** 06 April 2022; **Accepted:** 08 May 2022

Keywords: Cell classification; cell type; fluorescence microscopy; marker; segmentation

Abstract

Fluorescence microscopy techniques have experienced a substantial increase in the visualization and analysis of many biological processes in life science. We describe a semiautomated and versatile tool called Cell-TypeAnalyzer to avoid the time-consuming and biased manual classification of cells according to cell types. It consists of an open-source plugin for Fiji or ImageJ to detect and classify cells in 2D images. Our workflow consists of (a) image preprocessing actions, data spatial calibration, and region of interest for analysis; (b) segmentation to isolate cells from background (optionally including user-defined preprocessing steps helping the identification of cells); (c) extraction of features from each cell; (d) filters to select relevant cells; (e) definition of specific criteria to be included in the different cell types; (f) cell classification; and (g) flexible analysis of the results. Our software provides a modular and flexible strategy to perform cell classification through a wizard-like graphical user interface in which the user is intuitively guided through each step of the analysis. This procedure may be applied in batch mode to multiple microscopy files. Once the analysis is set up, it can be automatically and efficiently performed on many images. The plugin does not require any programming skill and can analyze cells in many different acquisition setups.

Impact Statement

Cell-type classification is an absolute requirement for quantitative analysis of microscopy imaging. Because different cell types normally differ in function and appearance, this tool allows researchers to correctly identify specific cells sharing common shape, life cycle, and phenotypical features. Here, we present Cell-TypeAnalyzer, a new plugin freely available under ImageJ or Fiji distribution for automated cell detection, identification, characterization, counting, and further cell-type classification based on user-defined criteria. Our tool aims at reducing the amount of subjectivity and human labor required to quantitatively assess the outcome of an imaging experiment.

1. Introduction

Nowadays, both multi-fluorescence imaging and labeling techniques are commonly used to identify biologically relevant processes through quantitative data extraction from fluorescently labeled molecules of interest[1–3]. Parallel to this unprecedented progress, advances in open-source bio-image software and scientific computing[4], cell counting automation, and single-particle analysis algorithms ensure reproducibility and objectivity compared to the more subjective manual analyses[5,6]. In cell biology, distinguishing specific cell types has traditionally been a labor-intensive and subjective task since it tries to

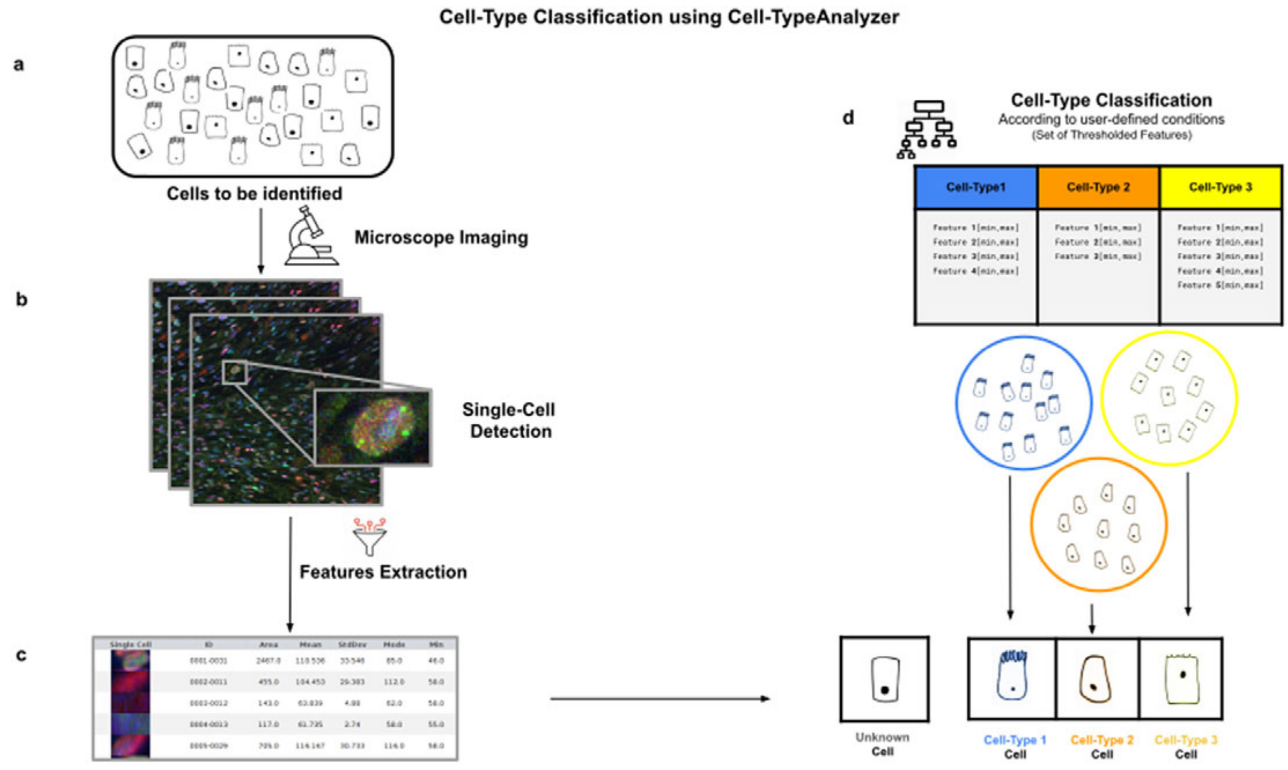

Figure 1. *Illustration of the workflow to identify specific cell types in a cell population. (a) Cell culture in which classification will be done to identify specific cell types. (b) Cell images are acquired and then processed for single-cell segmentation, feature extraction, and cell-type classification. (c) A collection of diverse features are extracted to both characterize and identify by ID number each cell. (d) Cell types are defined by a set of constraints in any of the detected features. The user may define as many cell types as needed, and each cell type is defined by as many constraints on the features as desired.*

classify cells according to morphological or phenotype forms[7] using tedious laboratory procedures as visual inspection. It is quite challenging to design a versatile algorithm to automatically identify different cell types on multiple fluorescent markers located on the same field[8,9] at the single-cell level. Additionally, in fluorescence microscopy, the signal-to-noise ratio is often low and the resolution quite limited[10], making automation of cell-type classification even more challenging[11]. In this context, single-cell features extraction arises helping researchers to properly overcome these drawbacks defining cell types which will be then cataloged into groups revealing different cell states or behaviors.

Cell-TypeAnalyzer allows the user to classify cells of interest (see Figure 1), identifying a set of cells sharing common morphological, intensity, or spatial features according to a given biologically defined class. Cell-TypeAnalyzer is an open-source plugin under the GNU public license that works equally well under Fiji[12] or ImageJ[13], offering a semiautomated cell-type classification using separated RGB channels for multiple microscopy image formats in an objective manner considerably more accurate than qualitative strategies. Therefore, Cell-TypeAnalyzer enables users describing a cell population through a set of extracted features to identify biologically relevant similarities or variations on a sample.

Our tool is highly configurable and may be adapted to many density or low-resolution situations by tuning the workflow internal parameters. We do not make any special assumptions about cell morphology, image formation process, optical microscope settings, or specimen features. In quantitative immunohistochemistry analysis, holding an accurate segmentation method to exactly isolate each cell from its possible fluctuating background is crucial to reach a robust detection[14,15]. In particular, in cases dealing with heterogeneous staining or overlapping cells, global auto-threshold methods may be a generic option to find the global optimal segmentation[16] grouping image pixels automatically depending on pixel values with no presumption about binary shapes or circularity, and hence leading to a less-biased detection non-exclusively limited to spot-like or roughly spherical objects[11]. Once the segmentation of each cell is

https://doi.org/10.1017/S2633903X22000058 Published online by Cambridge University Press

done, each one is measured as well as described individually by a vector of physical, morphological, statistical, and intensity features and then used for further cell-type classification.

Apart from that, this plugin was implemented under the ImageJ ecosystem to benefit from this bio-image platform mainly preferred and used by many life scientists. In recent years, researchers may choose from a wide range of open-source bio-image packages[17] to customize their own image analysis protocols through scripts, workflows, or plugin development. Nevertheless, many researchers may not have this computational proficiency. For such cases, Cell-TypeAnalyzer allows semi-automation based on a broadly applicable strategy for customized cell classification[18]. Furthermore, Cell-TypeAnalyzer is easily scriptable to customize the cell-type approach even in batch mode and obtain user-defined cell-type classifications across RGB channels dealing with multiple image formats currently supported by Bio-Formats[19]. Additionally, the user may choose a specific region of interest rather than considering the whole image. The researcher is guided through a user-friendly wizard-like graphical user interface (GUI) to perform each step. This GUI allows navigating forward or backward across panels to recalibrate settings in case of inadequate outputs.

2. Results

2.1. Overview of the procedure

Cell-TypeAnalyzer can work with images with up to three color channels. One of the channels, called Marker I, defines what a cell is and what is not. This channel can be a marker of cytoplasm, nuclei, or any other cellular structure of interest. Once we have identified cells with Marker I, cell types will be defined with Markers II and III.

A high-level overview of the Cell-TypeAnalyzer procedure involved is shown (see Figure 2). The processing actions consist of six major stages:

- Step I: After loading the raw RGB images, we need to establish the correspondence between the RGB channels and the marker names and roles. At this point, we may perform a spatial calibration (give the pixel size in physical units) to get measurements in real length units or pixels otherwise. We may also restrict the analysis to a region of interest which must be a closed shape. The plugin shows a histogram of the pixel values in each one of the RGB channels as visual feedback.
- Step II: The next step is the identification of the cells based on Marker I. To isolate cells from their background, we offer multiple possibilities. All of them respond to an auto-thresholding with different methods[20] to binarize the image, then a watershed transformation[21] may be applied to separate connected cells. Next, single-cell contours are detected and boundaries traced. Once done, features are extracted from each cell, and each cell obtains a unique ID number. The plugin shows at this point a summary of the detected features through some descriptive statistics (mean, median, variance, standard deviation, minimum, maximum, quantiles, inter-quantile range, etc.). The user may now apply filters based on these features to keep only the relevant cells for their study.
- Step III: Morphological operators[22] (erosion or dilation) may be applied to the cell contours to alter their original size. These operations allow the measurements on Marker II to be performed in a region that coincides with the area detected by Marker I (no operation), a smaller region (erosion), or a larger region (dilation). We may also perform a “Foci per nucleus”[23] analysis to count small bright dots within each cell. Then, we will compute different features of each cell from Marker II in the selected regions. These types of features are shape descriptors (to describe cell boundaries), shape metrics, and intensity-based statistics (calculated from intensity values in each channel on each cell). Finally, we may create cell types and, to each one, add as many constraints based on the Marker II features as needed.
- Step IV: We repeat the same actions as in Step III, but now on Marker III. Then, we can add the constraints on Marker III to the definition of each cell type. Cells are assigned to each one of the types

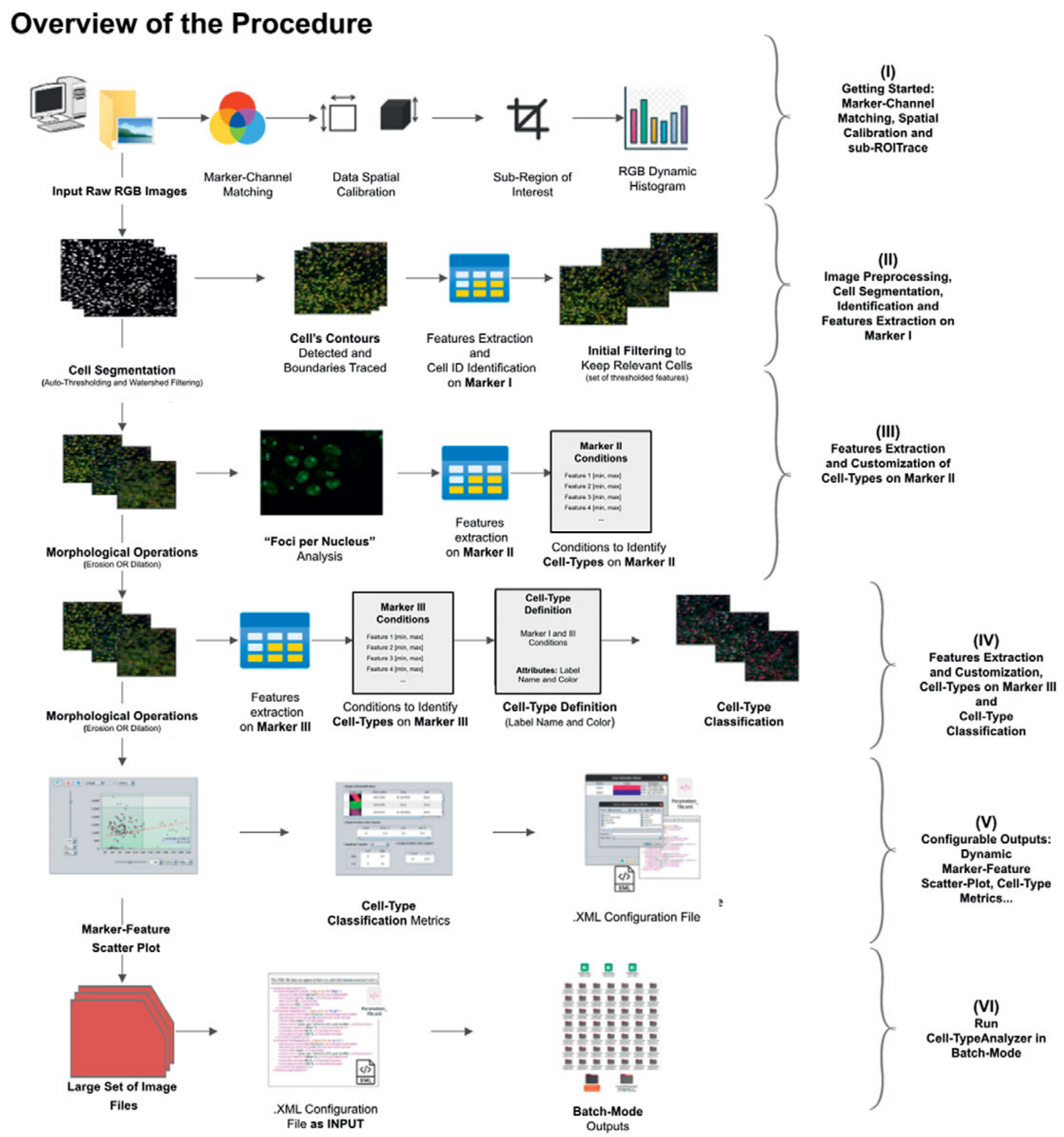

Figure 2. *Schematic overview of the Cell-TypeAnalyzer procedure to classify cells. (I) Marker-Channel Matching, data spatial calibration to have measurements on physical units, drawing a region of interest to restrict cell classification to a specific area. (II) Image preprocessing actions, cell segmentation (auto-thresholding and watershed transformation) to isolate cells from their background, identification by ID number, and feature extraction on Marker I. (III) Cell features are extracted on Marker II and declaration of the conditions of each cell type. (IV) Cell features are extracted on Marker III and modification of the cell-type conditions. (V) The user configures the output analysis. (VI) Cell-TypeAnalyzer is run in batchmode to large image sets.*

if they meet all the conditions on Markers II and III. Note that cell types can also involve conditions solely on Marker II or Marker III.

- Step V: The last interactive step allows us to configure a dynamic scatter plot to display any cell feature as a function of any other. Data points will represent relevant cells (those passing the criteria of a valid cell according to Marker I) being colored depending on their cell type or in gray if they do

https://doi.org/10.1017/S2633903X22000058 Published online by Cambridge University Press

not fulfill the criteria of any defined cell type. Finally, we may save an XML configuration file that will allow us to run this analysis in batch mode for many images (Step VI).

- Step VI: In this step, we apply the image analysis steps defined above (cell segmentation, region operations, etc.) and classify the detected cells into the user-defined cell types to a large number of images that have been acquired with similar characteristics as the one that served to set up the analysis. This execution is performed in batch mode and produces text or Excel files with the results for each image and a summary for the whole set.

In the following paragraphs, we go over each step in more detail.

2.2. Step I: Marker-channel matching, spatial calibration, and sub-ROI trace

Color images can be loaded, and a z-stack is automatically generated with each separated channel and the non-split RGB image. Thanks to the instant visualization, the user has efficient control over each operation performed. Through the "Channel Settings" panel, the user can establish the marker-channel matching defining those channels used for cell segmentation (Marker I) and further cell classification (Markers II and III). The "Calibration Settings" panel enables the user to have each feature spatially calibrated on physical units rather than pixels. The user must access the pixel size usually accessible within the image metadata to properly fill these spatial calibration fields. The "Crop Settings" panel allows defining a region of interest to be analyzed. Alternatively, the user may manually draw a closed area using any shape available on ImageJ's region-of-interest (ROI) tools[24]. If the crop option is employed, the X–Y coordinates of the detected cells are internally updated to reflect the current boundaries[25]. Finally, the last panel provides an overview of each marker's intensity pixel value distribution with a dynamic histogram. This workflow is schematically illustrated in Figure 3.

2.3. Step II: Image preprocessing, cell segmentation, identification, and feature extraction on Marker I

In this step, the user tunes the segmentation procedure to identify cells in the Marker I channel. In cases where the cell density is quite low or requires coping with a very noisy background, the user may apply some extra preprocessing actions to reduce noise using the preprocessing operations (image enhancement, correction, filtering, and de-noising) integrated by default in Cell-TypeAnalyzer. The summary of preprocessing methods is provided in Table 1. The user may apply as many preprocessing actions as needed by clicking on the Script button. A dialog window will pop up in which the user will be prompted by a script editor to write their own code in any of ImageJ's Macro supported language without saving or even, likewise, copying it to the clipboard and pasting it on the script editor area, then run it (Figure 4). Irrespective of using those preprocessing operations integrated by default or by scripting, these are applied to the image of Marker I, prior to cell segmentation.

The next step is the identification of the cells of interest in the Marker I channel. More than 10 global auto-thresholding algorithms (Default, Huang, Intermodes, IsoData, Li, MaxEntropy, Mean, MinError (I), Minimum, Moments, Otsu, Percentile, RenyiEntropy, Shanbhag, Triangle, and Yen) are available to binarize the image. The user should choose the one that best suits the specificities of the images being analyzed. It is common to find cells in close contact with other cells. Binarization algorithms cannot separate them into distinct entities. For this task, we provide a watershed segmentation[21] that works considering the output of the previous binarization. Finally, the cell contours are calculated, their boundaries traced, and features extracted (shape descriptors and intensity-based statistics) for each one of the cells are computed. To extract features from each cell, cell descriptors measure cell contours in the resulting binary image. These more than 20 different cell features computed are summarized in Tables 2 and 3.

To keep the relevant cells solely, the user may filter out irrelevant cells by defining thresholds in any calculated feature. These thresholds may be chosen with the help of scrolling sliders to define the minimum and maximum values of any feature to be considered a relevant cell. The cell-type classification

https://doi.org/10.1017/S2633903X22000058 Published online by Cambridge University Press

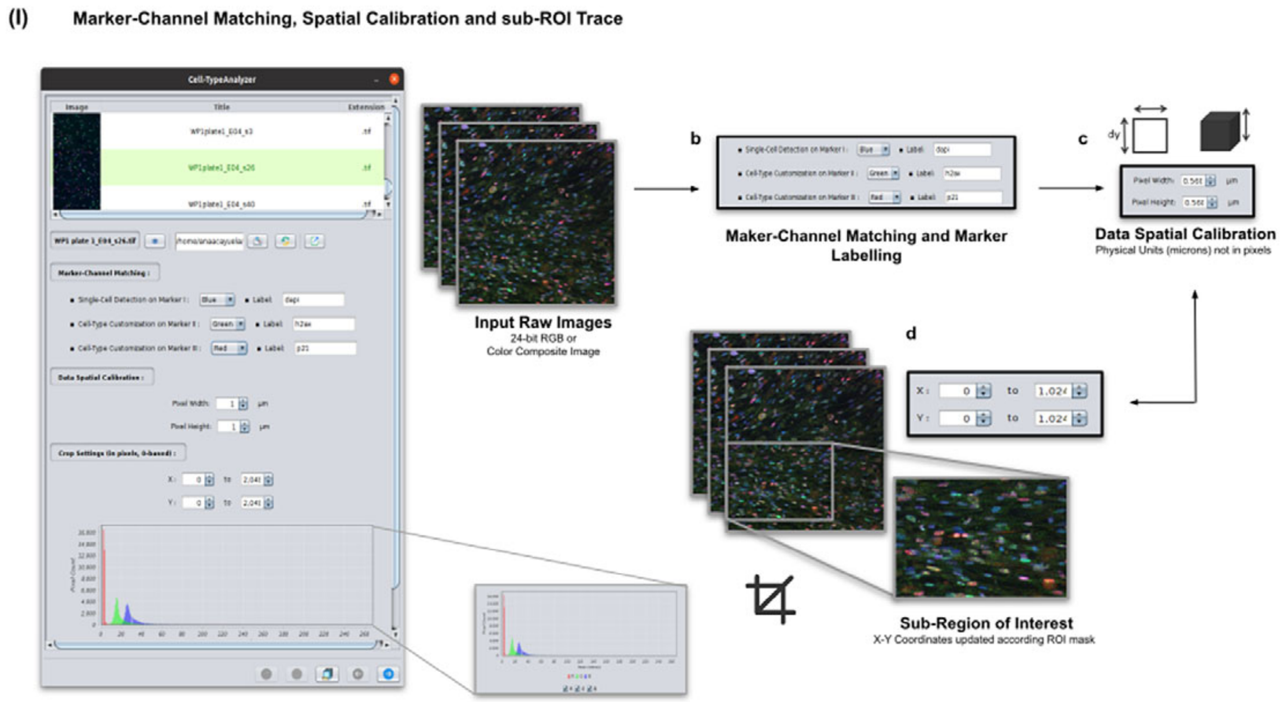

Figure 3. *Details of Step I. Marker-channel matching, spatial calibration, and sub-ROI trace workflow. (a) Images to be processed must be in 2D single-plane RGB form: 24-bit RGB or Color Composite. (b) Via "Marker-Channel Matching," the user must determine the matching between the RGB channels and the Markers I–III. (c) Through the "Data Spatial Calibration" panel, the user obtains all cell metrics calibrated on physical units (not in pixels) by typing the image pixel size. (d) The "Crop Settings" panel enables to draw a region of interest to be considered for analysis. All coordinates calculated throughout the plugin are updated according to the location of the closed shape. In addition, the user may inspect, by clicking on a dynamic histogram, the distribution of pixel intensities on each marker.*

of the subsequent steps is only applied to those cells that have been flagged as relevant in this cell identification step based on Marker I.

2.4. Step III: Features extraction and customization of cell types on Marker II

Once cells have been successfully segmented on Marker I, the features extraction of relevant cells on Marker II is performed. This stage may involve morphological operations[22] such as erosion or dilation (see Figure 5a) or, conversely, none to maintain the original cell area. If required, the user can perform a "Foci per nucleus" analysis[23] to count all small bright dots (local maxima of pixel intensity)[20] within each cell (see Figure 5b). This analysis has some tunable parameters like the "Tolerance" (by default 30), which acts as a local threshold (a maximum is removed from the list if it is close to another one within a distance smaller than "Tolerance").

Features of Marker II are finally calculated for the relevant cells. This information is attached to the information already known for each cell after the analysis of Marker I. As for Marker I, this information is displayed in a data table, and a statistical summary is shown. At this point, the user may define cell types creating as many constraints as needed for the features computed on Marker II (e.g., a cell is of Type 1 if its area is between this and this value, its circularity between this and this, etc.).

2.5. Step IV: Features extraction and customization of cell types on Marker III

This step is totally analogous to the previous one on Marker II. The definition of cell types can include conditions on any feature of both Markers II and III. A cell is classified in these types if it fulfills all the

Table 1. *Table listing the preprocessing operations which may be applied by default using Cell-TypeAnalyzer previous to image thresholding.*

Summary of preprocessing actions		
Action	Type of action	Description
Smooth	Filter	Blurs the image replacing each pixel with the average of its 3×3 neighborhood.
Sharpen	Filter	Increases contrast and accentuates detail in the image replacing each pixel with a weighted average of the 3×3 neighborhood.
Enhance contrast	Contrast adjuster	Enhances image contrast by using either histogram stretching or histogram equalization.
Gaussian blur	Filter	Uses convolution with a Gaussian function for smoothing.
Median	Filter	Reduces image noise by replacing each pixel with the median of the neighboring pixel values.
Mean	Filter	Smooths image by replacing each pixel with the neighborhood mean.
Unsharp mask	Filter	Subtracts a blurred image copy and rescales the image to obtain the same contrast of large structures as in the input image.
Minimum	Filter	Grayscale erosion replacing each pixel in the image with the smallest pixel value in that pixel's neighborhood.
Maximum	Filter	Grayscale dilation replacing each pixel in the image with the largest pixel value in that pixel's neighborhood.
Variance	Filter	Highlights image edges by replacing each pixel with the neighborhood variance.

conditions of that cell type. Cells that do not fulfill any defined cell types are classified as unknown (see Figure 6).

2.6. Step V: Configurable outputs, dynamic marker-feature scatter plot, and cell-type metrics

At this stage, the user can dynamically configure a 2D analysis (see Figure 7) to plot any cell feature extracted from the selected marker (Marker I, II, or III) as a function of any other. This functionality may be beneficial to observe relationships among features from markers. Users may choose to apply different curve-fitting models (Linear, Polynomial, Power, Logarithmic, or Exponential) to find the one that best fits data (see Figure 7c). The 2D analysis may be restricted to specific z-slices. This is useful to identify possible dependencies on the cell's height within the tissue in confocal microscopy (e.g., apical vs. luminal cells). Each point represents a relevant cell. Its label determines its color on the plot. This helps to recognize cell types' distribution patterns according to their location in specific feature planes (see Figure 7a). Cells not belonging to any cell type are colored in gray. As an additional way to explore the set of cells, the user may define thresholds for each plotted feature (see Figure 7a). These thresholds divide the feature space into four quadrants, and the number of cells in each quadrant is counted and displayed as a table.

If the user is satisfied with the analysis performed on several input images, the whole analysis description can be saved as an XML file used by the plugin in batch mode (see the next section).

2.7. Step VI: Running Cell-TypeAnalyzer in batch mode

To achieve the most accurate batch-mode analysis (see Figure 8), it is advised to perform prior tests to find the most suitable parameters on a subset of images before applying it to a large batch. The batch-mode

https://doi.org/10.1017/S2633903X22000058 Published online by Cambridge University Press

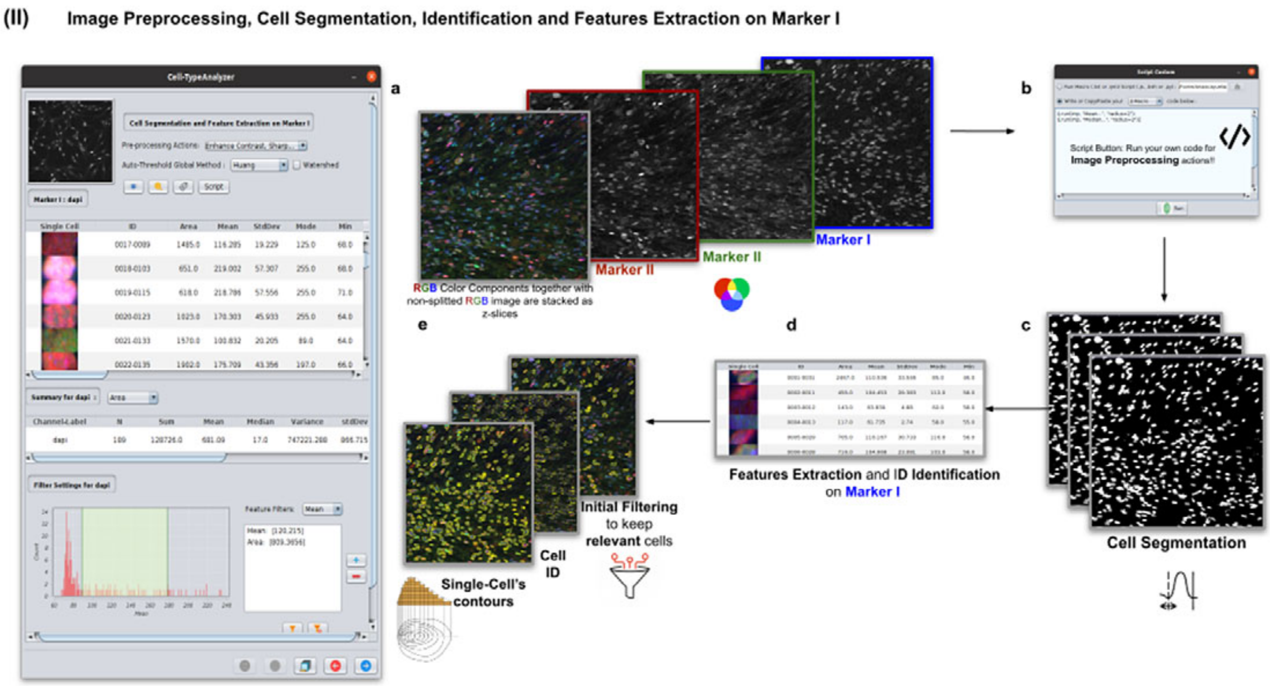

Figure 4. *Details of Step II. Image preprocessing, cell segmentation, identification, and feature extraction on Marker I. (a) The channel corresponding to Marker I is separated from the rest of the images. (b) Now, the parameters for cell segmentation on Marker I are tuned. The user may choose from different global auto-threshold algorithms to binarize the image and isolate cells from their background. In the event of having some connected cells, watershed filtering may be applied to split touching objects. (d) Features extraction from each cell. (c) Optionally, the user may provide an image preprocessing script that facilitates the identification of the cells of interest. This is done by clicking on the "Script" button and selecting a script file or writing their own code in any of ImageJ's supported languages. (e) Filtering to keep only relevant cells through scrolling sliders.*

GUI will ask the user to load the XML configuration file containing the whole list of user-defined processing actions. This file is saved in Step V. Using XML is advantageous because it is editable. Thus, the user can easily change the analysis without reopening the GUI and designing the analysis from scratch. Then, the user must supply the directory with the input images and the directory for the outputs.

3. Experimental Validation

To validate Cell-TypeAnalyzer and demonstrate how versatile it is in solving specific biological problems in several image sets, we propose different applications in which this tool may be customized.

3.1. Comparing cell-quantification using confocal and widefield microscopy

In the first application, we compared the count of cells of a given type using Confocal or Widefield microscopy[26]. We did not expect a significant difference between the two kinds of microscopy despite their different appearances in this experiment. We had two different cell preparations: control and treated cells (see Figure 9). A larger growth rate characterized the control group, and microscopy fields showed a higher cell density, while the treated group had a lower cell density[27]. Images in each well were acquired containing channels for 4',6-Diamidino-2-Phenylindole (double stranded DNA staining) (DAPI) as Marker I; this marker is a dye for targeting the cell nuclei[28]. As Marker II, we used Rabbit antihuman p21 (at 1:500 dilution), and antibody labeled cells were visualized with Goat anti-rabbit secondary antibody directly conjugated to fluorochrome Alexa 647 (at 1:500 dilution). As Marker III, we used Mouse antihuman phospho-histone H2AX (at 1:500 dilution), a biomarker to

https://doi.org/10.1017/S2633903X22000058 Published online by Cambridge University Press

Table 2. *Table reporting the types of features (shape descriptors and intensity-based statistics) computed for each cell along with description using Cell-TypeAnalyzer.*

Cell features summary		
Feature	Type of feature	Description
ID	ID number	Identification number of each cell contour
Area	Shape metric	Area of cell contour in squared pixels or physical units (if calibration is done)
Mean gray value (Mean)	Intensity-based statistics	Average gray value of all pixels within the cell contour
Standard deviation (StdDev)	Intensity-based statistics	Standard deviation of the gray values used to generate the mean gray value
Modal gray value (Mode)	Intensity-based statistics	Most frequent gray value within each cell contour
Min and max gray levels (Min & Max)	Intensity-based statistics	Minimum and maximum gray values within cell contour
X–Y centroid (X & Y)	Shape metric	Mean of the x and y coordinates of all pixels within each cell contour
X–Y Center of mass (XM & YM)	Shape metric	Brightness-weighted mean of the x and y coordinates of all pixels within cell contour
Perimeter (Perim.)	Shape metric	Length of each cell-contour boundary
Bounding rectangle (BX & BY, Width & Height)	Shape metric	Smallest rectangle enclosing each cell contour defined by "BX" and "BY" (the coordinates of upper left corner), and "Width" and "Height"

recognize DNA damage[29], being visualized with Goat anti-mouse secondary antibody directly conjugated to fluorochrome Alexa 488. Both Confocal multispectral system Leica STELLARIS 5 system and Leica DMi8 S Widefield epifluorescence were employed for the image acquisition. Each field of view was operated at a map resolution format of 1,024 × 1,024 pixels with each channel at 16-bit intensity resolution. The confocal images were acquired with an HC PL APO CS2 20 × /0.75 DRY objective and have a pixel size of 0.758 × 0.758 microns. The widefield images were acquired with an HC PL FLUOTAR L 20 × /0.40 DRY objective and have a pixel size of 0.65 × 0.65 microns. Images were acquired with a step size of 2.5 microns and intervals of approximately 9 s per image (44 sites per well), resulting in barely 10 min per well. A total of 128 wells were imaged using three channels, which resulted in 384 grayscale images. A dataset of 32 images per group (confocal-control, confocal-treated, widefield-control, and widefield-treated) was collected.

For the image analysis, we used Otsu's binarization to identify cell nuclei in Marker I. We did not need to use watershed segmentation to separate nearby cells. We removed all cells whose nucleus had a Marker I area below 5 pixels. We performed a "Foci per nucleus" analysis on Marker II. We defined a single cell type by requiring cells to have an average intensity in Markers II and III above the average intensity in Marker I and with at least eight foci. The analysis time per image was 6 s in a laptop Alienware M15 8th Gen Intel Core i7-8750H (4.1 GHz). The results of Cell-TypeAnalyzer in batch mode are shown in Figure 9. In Table 4, we compare the mean (two-sample t-test) between the confocal and widefield imaging. None of the tests could be rejected at a confidence level of 95%. This comparison shows that imaging with confocal or widefield microscopy does not make any statistically significant difference for this experiment. Cell-TypeAnalyzer was instrumental in automating this comparison, which would have been much more tedious if manual counting was required.

Table 3. *Continuation of table reporting the types of features (shape descriptors and intensity-based statistics) computed for each cell along with description using Cell-TypeAnalyzer.*

Cell features summary (continued)		
Feature	Type of feature	Description
Fit ellipse (Major, Minor, Angle)	Shape metric	Fit an ellipse to the cell contour being "Major", "Minor" the primary and secondary axis, and "Angle" the angle among primary axis and line parallel to the x-axis of cell contour
Circularity (Circ.)	Shape descriptor	Being a value of 1.0 a perfect circle and a value of 0.0 an elongated shape
Aspect ratio (AR)	Shape descriptor	"Major" divided by "Minor" axis
Roundness (Round)	Shape descriptor	Inverse of "Aspect ratio"
Solidity	Shape descriptor	"Area" divided by convex Area
Feret's diameter (Feret, FeretAngle, MinFeret, FeretX, and FeretY)	Shape metric	"FerretAngle" angle among Ferret's diameter and parallel line to the cell contour's x-axis; "MinFerret" is the minimum caliper diameter; "FerretX" and "FerretY" the starting coordinates of Ferret's diameter
Integrated density (IntDen, RawIntDen)	Intensity-based statistics	Product of "Area" and "Mean Gray Value"
Median	Intensity-based statistics	Median value of pixels within each cell contour
Skewness (Skew)	Shape metric	Third-order moment about the mean
Kurtosis (Kurt)	Shape metric	Fourth-order moment about the mean
Area fraction (%Area)	Intensity-based statistics	Percentage of nonzero pixels

3.2. Evaluating morphology of known phenotypes in HeLa cells

This second application proposes a general approach to identify and get subsequent phenotype classification of single cells within cell populations based on morphological changes.

This study used the Cell-TypeAnalyzer plugin in batch mode to efficiently extract shape descriptors and features from cells to analyze their morphology and obtain fluorescence statistics from nuclear and cytoskeletal markers. Each cell was described by a vector of more than 30 descriptors measured for each fluorescent marker (DNA, Tubulin, and Actin).

Using this tool, we established a semiautomated method for identifying distinct phenotypes from subsequent classification, according to classes that are user-defined for each cell type. These classes consist of parameters based on morphology and cell area to describe their protrusion or elongation. HeLa cells in this application were first detected, identified by an ID number, and finally, classified into different cell types: Actin fiber (AF), Big cells (BC), Condensed cells (C), Metaphase cells (M), Normal cells (N), and Protruded cells (P). The full dataset of HeLa cells was downloaded from the image data resource (IDR)[30], a public repository of high-quality bio-image datasets from published scientific studies. Specifically, images were selected from the "idr0012-fuchs-cellmorph" dataset[31]. This dataset consists of 22,839 siRNA-mediated knockdowns on HeLa cells in which genes were clustered, and their function predicted on a genome-wide scale. The workflow for cell-type classification (see Figure 10) starts with the "Splitting multi-channel images" command, which is called automatically by Cell-TypeAnalyzer. This command is used for color image processing since it splits the RGB images into

https://doi.org/10.1017/S2633903X22000058 Published online by Cambridge University Press

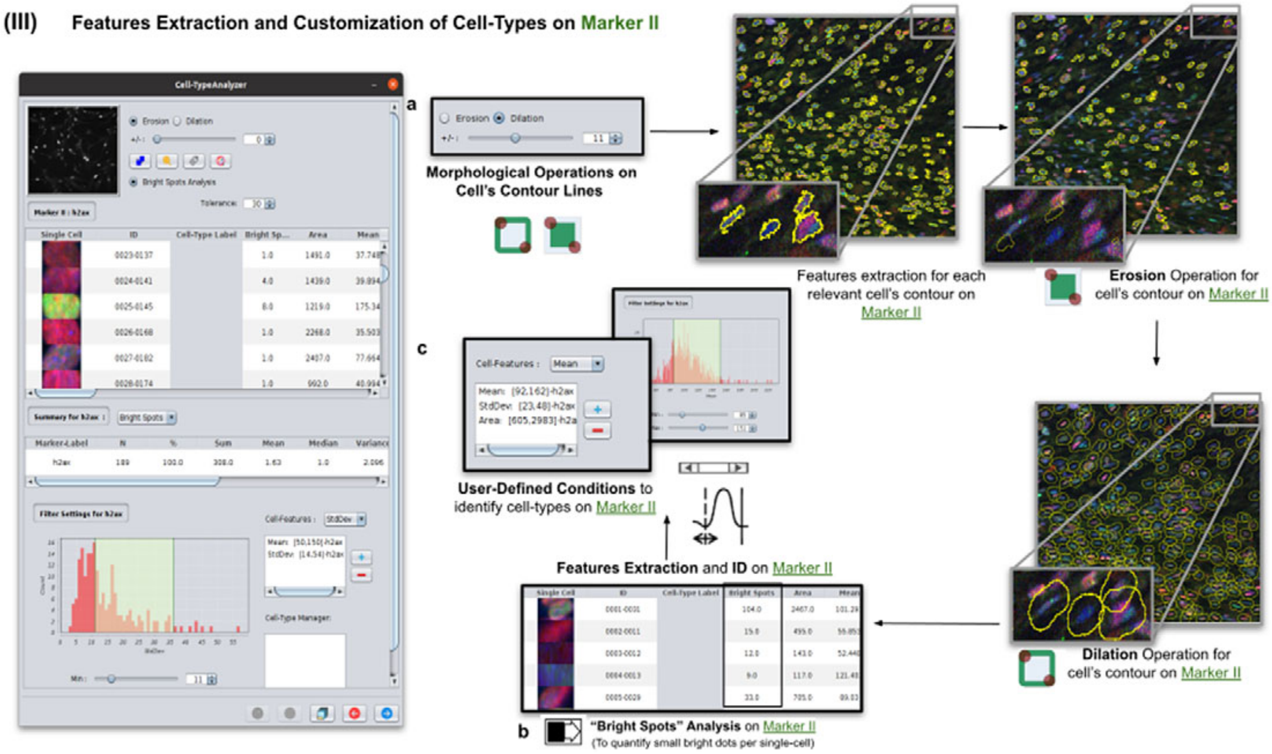

Figure 5. *Details of Step III. Features extraction and customization of cell types on Marker II. (a) Considering as reference the relevant cells, the user may apply morphological operations (erosion and dilation) on cell contours to resize them. (b) A "Foci per nucleus" analysis[23] may be performed whose goal is to quantify the small bright dots within each cell contour. Finally, the features extraction of relevant cells is done on Marker II, attaching this vector to the description of each relevant cell. (c) The user defines cell types based on values of the features calculated on Marker II.*

their respective nuclear (DNA) and cytoskeletal (Actin and Tubulin) components. Nuclei and cytoplasm boundaries were isolated through segmentation using the "Auto-Threshold" Otsu's and Huang's methods, respectively. The watershed segmentation was applied to separate touching nuclear or cytoskeletal structures. Once these regions were isolated, features were measured from both cytoskeletal fluorescent markers (Tubulin and Actin) for each cell. An initial filtering was applied to remove those regions having an area in pixels smaller than 20. The remaining cells were classified depending on their quantified fluorescent intensity on the Actin marker and their respective circularity values. Those cells having more intensity in the Tubulin marker than the Actin marker and a circularity value located in the Q4 quartile of the distribution for circularity were classified as Metaphase (M) cells. Otherwise, cells were classified as Actin fibers (AF) class. Cells were classified into the Big cells (BC) class if their area belonged to the Q4 quartile. The remaining cells were considered as candidates to belong to the Normal (N), Condensed (C), or Protruded (P) cell type. This classification was performed as a function of the circularity: Protruded (if the circularity was in the Q1 quartile), Normal (Q2 or Q3), and Condensed (Q4). The proportions of cells in each one of the types are similar to the one originally reported in Reference (31).

3.3. Classifying morphology in Spirochaete bacteria on dark-field microscopy

This third application proposes a widely applicable analysis workflow to detect, identify by ID, quantify, and get subsequent morphological phenotyping of *Spirochaete* bacteria in blood. We applied our tool on all dataset images in Reference (32), a total of 366 dark-field microscopy images. It must be noted that these images are monochromatic, showing that our tool is not restricted to the analysis of multichannel images. The phenotype classification was done using the Cell-TypeAnalyzer plugin in batch mode. Cell

https://doi.org/10.1017/S2633903X22000058 Published online by Cambridge University Press

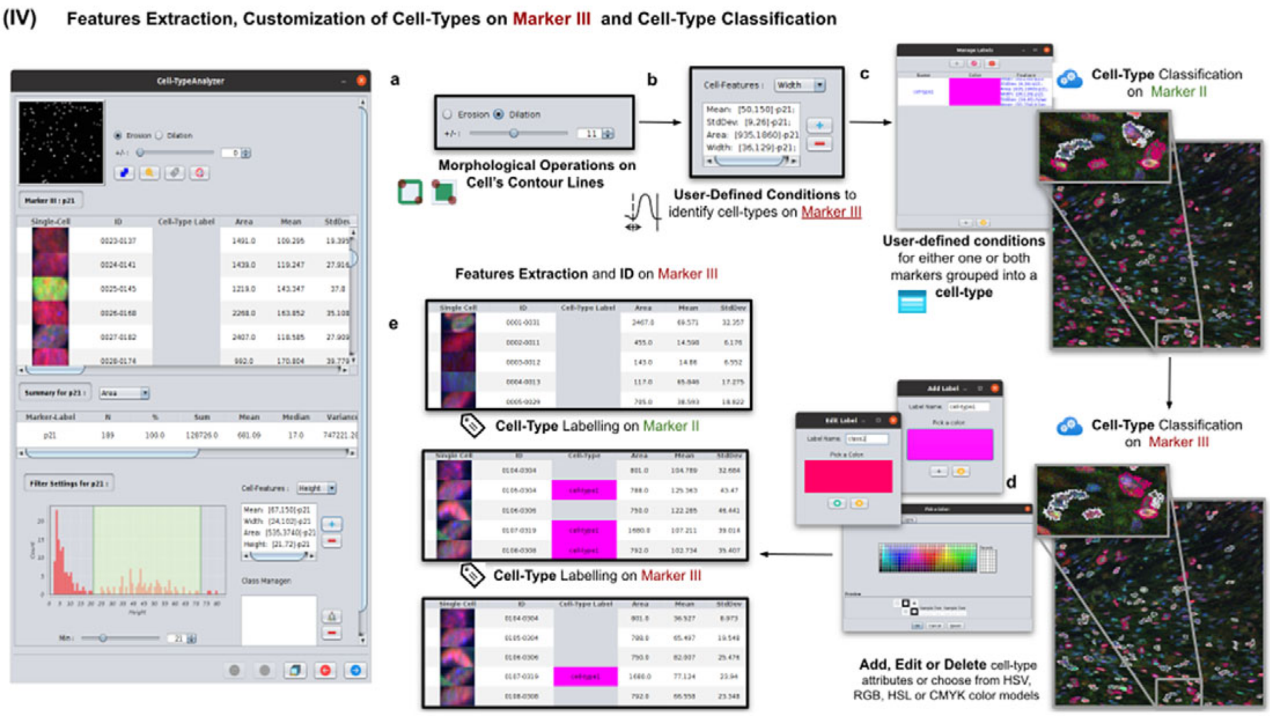

Figure 6. *Details of Step IV. Features extraction and customization of cell types on Marker III. (a) User may either erode or dilate cell-contour lines. Then, features are extracted from relevant cells on Marker III, generating, once more, a vector for each cell. (b) Once done, the user may define conditions on any of the Marker III features to refine the definition of cell types further. (c) Cell-type labeling and coloring can be defined by the user. (d) Cell-type conditions can be iteratively defined between Steps III and IV until the desired labeling is achieved. (e) For each detected cell, a label is attached depending on which conditions it fulfills. This operation helps to refine the definition of the cell types.*

features (shape descriptors and intensity-based statistics) were automatically extracted from every cell as a vector used to classify them into different morphological classes.

The evaluation of bacterial single-cell contours may be instrumental in getting new insights into the morphological changes due to a wide range of processes that external perturbations may induce and thus reflected in the cell shape. Although *Spirochaete* organisms normally show stable, well-defined shapes, these single-cell microbes may change their morphology in response to certain environmental signals or even, depending on their life-cycle stage. The following morphological cell types were defined: Blood Cells (BC), Normal (N), Small (S), Elongated (E), and Round (R). In this analysis, around 53,000 cells were automatically segmented, identified, and classified into each cell-type class.

The workflow for phenotype classification (see Figure 11) on bacterial cells starts with the “Splitting multi-channel images” command to separate the RGB images to their respective color components. Blood cells and *Spirochaete* bacteria boundaries were isolated through segmentation using the “Auto-Threshold” Otsu’s method to binarize the image. Then, the watershed segmentation was applied. The “Fill Holes” command was called obtaining more homogeneous regions. Once these regions of interest were isolated, cell features described in Section 5 were calculated for every cell. Candidate cells were first classified as blood cells or bacteria cells depending on their area in pixels. Subsequently, bacteria cells with circularity values located in the Q4 quartile were classified as Round (R) cells. The remaining cells were classified as Elongated (E), Small (S), or Normal (N) according to their area. Hence, bacteria cells showing an area in pixels located in the Q4 quartile were labeled as Elongated €, then those having area values belonging to Q1 were identified as Small (S), and finally, those located both at inter-quartile range were classified as Normal (N).

https://doi.org/10.1017/S2633903X22000058 Published online by Cambridge University Press

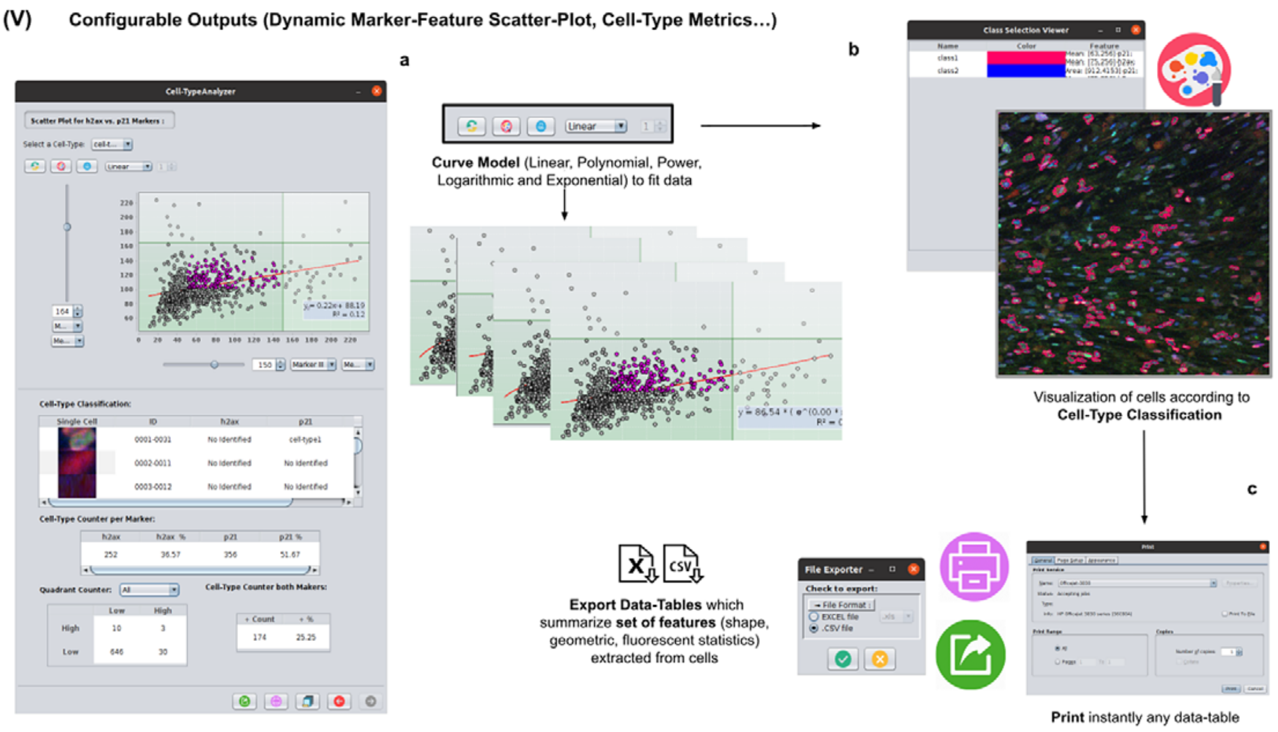

Figure 7. *Details of Step V. Configure outputs (dynamic marker-feature scatter plot and cell-type metrics). (a) The user can dynamically plot any cell feature from either marker (Marker I, II, or III) as a function of any other. Different curve models (linear, power, polynomial, and logarithmic) can fit the data. A point represents each cell. If the cell is classified under a specific cell type, the corresponding cell-type color is used. Otherwise, cells that do not belong to any cell type are colored in gray. (b) Contours from cells belonging to a specific cell type may be visualized as outlines. (c,d) There are multiple ways of exporting the analysis, including CSV, Excel files, and PDF prints.*

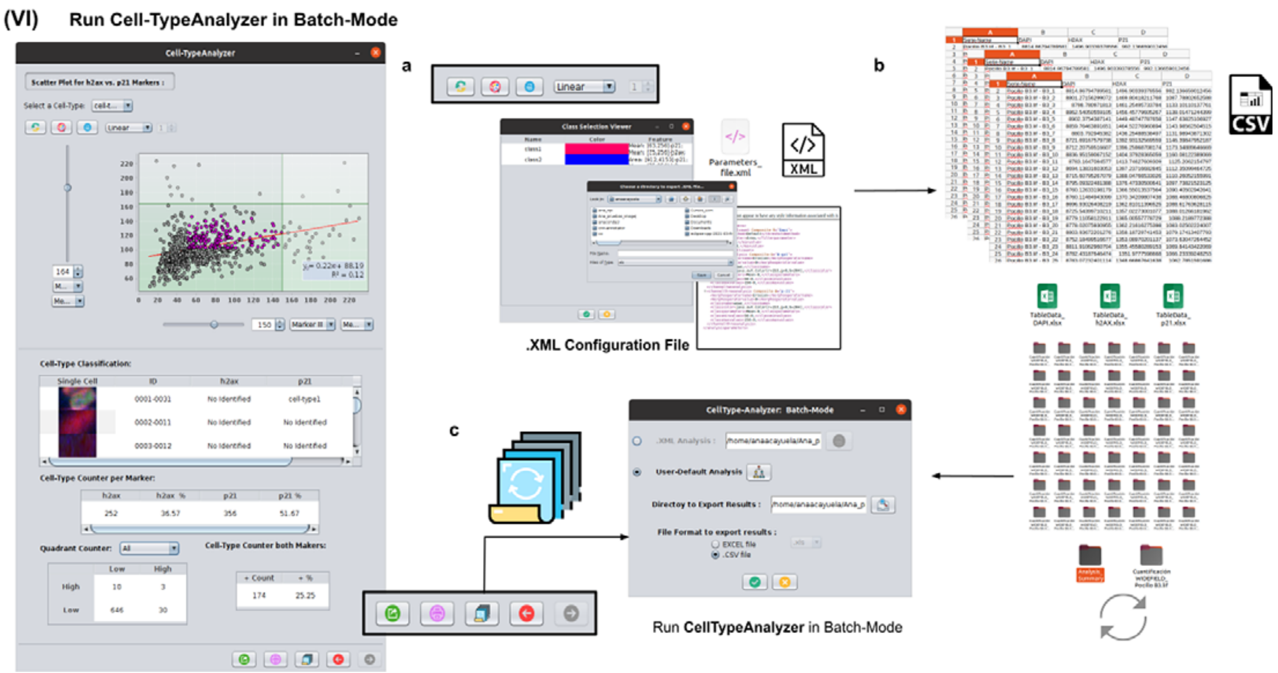

Figure 8. *Details of Step VI. Execution of Cell-TypeAnalyzer in batch mode. (a) The user may save an XML configuration file that summarizes all the steps required, and it will be used to run Cell-Type-Analyzer for large sets of images. (b) Examples of output files generated. (c) Graphical user interface for batch mode.*

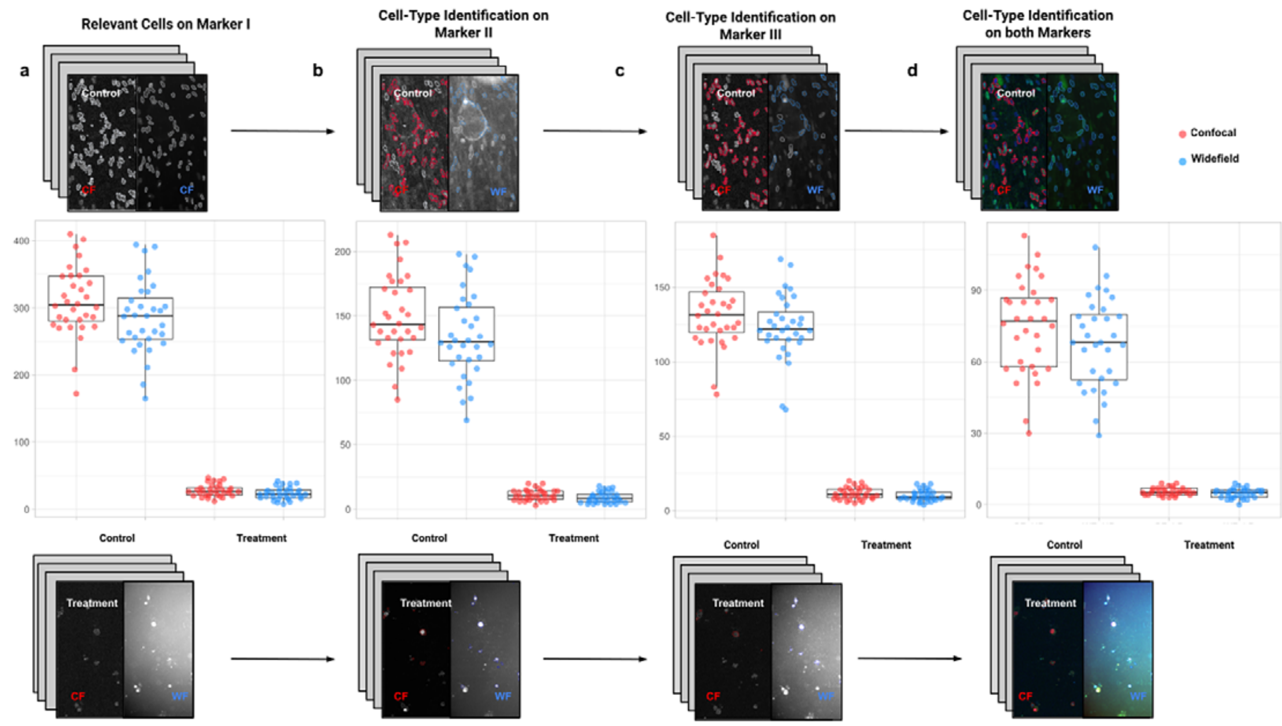

Figure 9. *Box-whisker plots summarizing the distribution of both control and treatment groups from confocal and widefield microscopes values. Each point will be representing the total number of cells quantified for each analyzed well. (a) Data points calculated by quantifying relevant detections for control and treatment samples on Marker I (DAPI). The distribution charts reveal nonsignificant differences between microscopes. (b,c) Data points calculated by quantifying cells identified within a specific cell type on Markers II and II, respectively. The distribution reveals nonsignificant differences among microscope tested. (d) Data points calculated by quantifying cells that are identified simultaneously as a specific cell type for both markers.*

4. Discussion

Cell-TypeAnalyzer was developed to facilitate researchers the single-cell identification and subsequent cell-type classification under user-defined conditions. Furthermore, Cell-TypeAnalyzer may measure large sets of images with many cell types of interest previously defined by the user in a large variety of biological samples, an aspect which is increasingly recognized as crucial to improve our understanding of how genetic and environmental factors give rise to changes in organisms or even in their behavior[33].

Two of the most popular image analysis software for identifying and quantifying cell phenotype are CellProfiler[34] and CellProfiler Analyst (CPA)[35]. CellProfiler is a flexible and open-source image analysis software package which allows users to mix and match modules to create their own customized image analysis pipelines without extensive programming skills. CPA was released in 2008 and marked a great progress in fully automated phenotypic analysis including modern statistical learning methods to identify specific cell phenotypes. This software (directly interfaced to CellProfiler) enabled biologists to define a bunch of phenotypes as well as create annotations for single cells to train supervised machine learning algorithm to further predict phenotypes on unseen data. Notwithstanding the first release of CellProfiler exhibited several constraints in terms of the definition of classes (only supported two classes: positive and negative), and on its behalf, CPA provided a small number of machine learning algorithm available for classification (only GentleBoost), both tools were extensively used worldwide. Nowadays, those limitations have been amply overcome since recent releases have incorporated the definition of multiple phenotype classes as well as different machine learning algorithms are currently supported.

Having shown that trends in phenotypic analysis go through using platforms in fully automated mode, we would like to make a comparison with our tool for cell-type classification functioning in

Table 4. *Table showing descriptive statistics for both cell populations (Confocal and Widefield) depending on Control or Treatment conditions. Panel A: Control—means and distributions are nonsignificantly different between Widefield and Confocal microscopes at the 0.05 level in* t*-test. Panel B: Treatment—means and distributions are nonsignificantly different between Widefield and Confocal microscopes at the 95% confidence level in* t*-test. Since* p*-value > .05, the average of WF's population cannot be rejected from being equal to the average of the CF's population.*

	Marker I		Marker II		Marker III		Cell types	
	WF[a]	CF[b]	WF	CF	WF	CF	WF	CF
Panel A: Control								
Mean	285.59	309.63	134.25	149.56	123.41	132.16	67.84	74.72
SD	54.17	51.9	32.99	31.95	21.88	22.62	18.64	19.88
p-value	0.075		0.064		0.121		0.159	
Effect size	0.45		0.47		0.39		0.36	
r[c]	0.978		0.983		0.973		0.987	
Panel B: Treatment								
Mean	23.47	27.91	9.03	11	10.06	11.41	4.69	5.53
SD	8.97	9.06	4.26	4.24	3.76	4.02	2.05	1.67
p-value	0.053		0.069		0.172		0.076	
Effect size	0.49		0.46		0.35		0.45	
r[c]	0.989		0.933		0.885		0.814	

Notes: Statistical significance depending on p-value at the $p < .05$ level. p-values were determined by using two-sample t-test for expected difference between two populations' mean ($n = 32$).
[a]Widefield microscope technique.
[b]Confocal microscope technique.
[c]Pearson correlation coefficient.

semiautomatic mode. All of these widely used tools which perform phenotypic analysis on images of cell-based assays are designed to operate fully automatically. This aspect is absolutely great in terms of processing speed and objectiveness, but sometimes might compromise accuracy regarding the user interpretation of extracted features which are prone to unexpected errors in case of users not having extensive experience. Moreover, this lack of interpretation may be more obvious as regards the complex structure of deep learning neural networks as well as the sophisticated of machine learning models that regularly require a prior knowledge of the field to be applied. On the other hand, semiautomated tools always require user input and interaction along with expert validation to extract the required information accurately and these methods, normally, are quite dependent on the quality of raw image data. Thereby such manual intervention might be certainly time-consuming, and it may introduce subjective bias leading to hind the implementation of these approaches to large sets of images.

In this context, there is a general growing need of universal tools for varying image conditions to accomplish semiautomated phenotypic analysis that enable cell-type classification within ImageJ or Fiji ecosystem. The development of Cell-TypeAnalyzer within ImageJ might help to remove this bottleneck in experimental pipelines which often involve complex workflows for a non-experienced community, offering a user-friendly solution relying less on fully automation. Moreover, Cell-TypeAnalyzer may be a worthy contribution with many options to customize cell-type analysis on multi-fluorescent microscopy images containing hundreds of objects, but it might be broadly applicable to other heterogeneous microscopy samples.

Consequently, an important difference between the more advanced software described above and our approach is that the advanced tools do not allow the user to fine-tune the parameters used for the classification in an understandable way, making outputs challenging to verify by user. For that reason, we

https://doi.org/10.1017/S2633903X22000058 Published online by Cambridge University Press

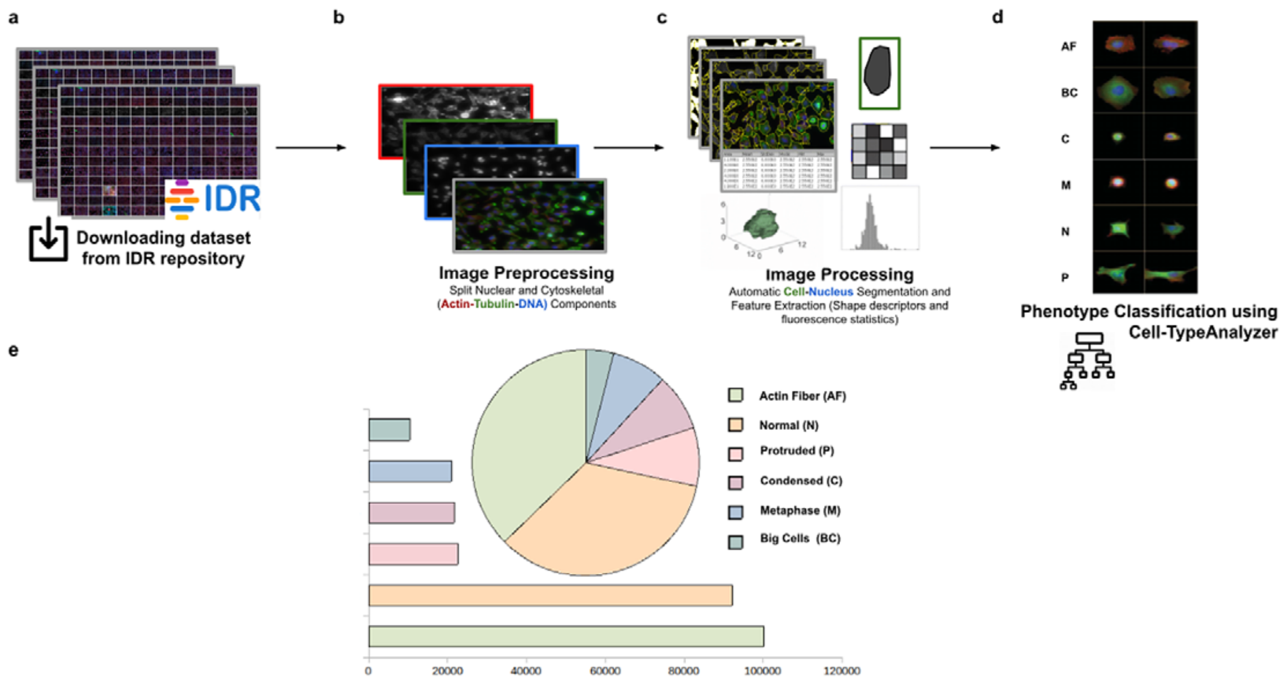

Figure 10. *Scheme of semiautomated analysis of raw images for classifying cellular phenotypes in HeLa cells. (a) The full dataset of HeLa cells images to be analyzed was downloaded from the image data resource repository. (b) Image preprocessing actions to get the separated nuclear (DNA) and cytoskeletal (Actin and Tubulin) components were applied. (c) Image processing actions for Cell–Nucleus segmentation and subsequent identification. A vector describes each cell based on shape descriptors, geometry, and fluorescence statistics. (d) Cells were classified into Actin Fiber (AF), Big cells (BC), Condensed (C), Metaphase (M), Normal (N), and Protruded (P) cell-type classes depending on user-defined feature conditions set for each case. (e) Quantification results of classifying HeLa cells belonging to each cellular phenotype.*

have designed Cell-TypeAnalyzer, that is modularly designed through a simple wizard-like GUI to visually guide user to each step of the analysis offering an instant visualization of the outputs for each marker, hence enabling manual verification across the navigation back and forward through wizards. As is already the case of CPA, which has an interactive GUI to view images, Cell-TypeAnalyzer allows user to manually scroll through a gallery view of image thumbnails corresponding to the input folder in which samples are located to directly being chosen by user for being analyzed avoiding the tedious of browsing directories. Additionally, Cell-TypeAnalyzer benefits from ImageJ ecosystem, which is probably the best-known, flexible, and longest-lived software for biomedical sciences and beyond. In consequence, Cell-TypeAnalyzer leverages from a lot of plugins for scientific image processing included within its distribution, such as Bio-Formats library, which deals with more than 150 different file formats. Even though CellProfiler offers quite powerful tools for detecting, quantifying, and describing cell morphology, Cell-TypeAnalyzer benefits from a large library of tools within Fiji distribution such as MorphoLibJ[22] for morphological filtering as well as reconstruction and global ImageJ thresholding for binarization/segmentation. Additionally, more experienced users may develop their own macro programs to automate image preprocessing actions such as brightness correction, pixel and geometric transformations, or even image filtering and restoration using ImageJ macros action integrated within Cell-TypeAnalyzer plugin. At this point, it must be noted that Cell-TypeAnalyzer was not developed to remove noise or enhance quality from images, although it provides preprocessing tools to improve cell detection and cell-type characterization. Best practices for acquisition are, whenever possible, recommended before using this tool. As is generally known, ImageJ was traditionally designed for single-image processing. On the contrary, CellProfiler was originally devised for building large-scale and modular

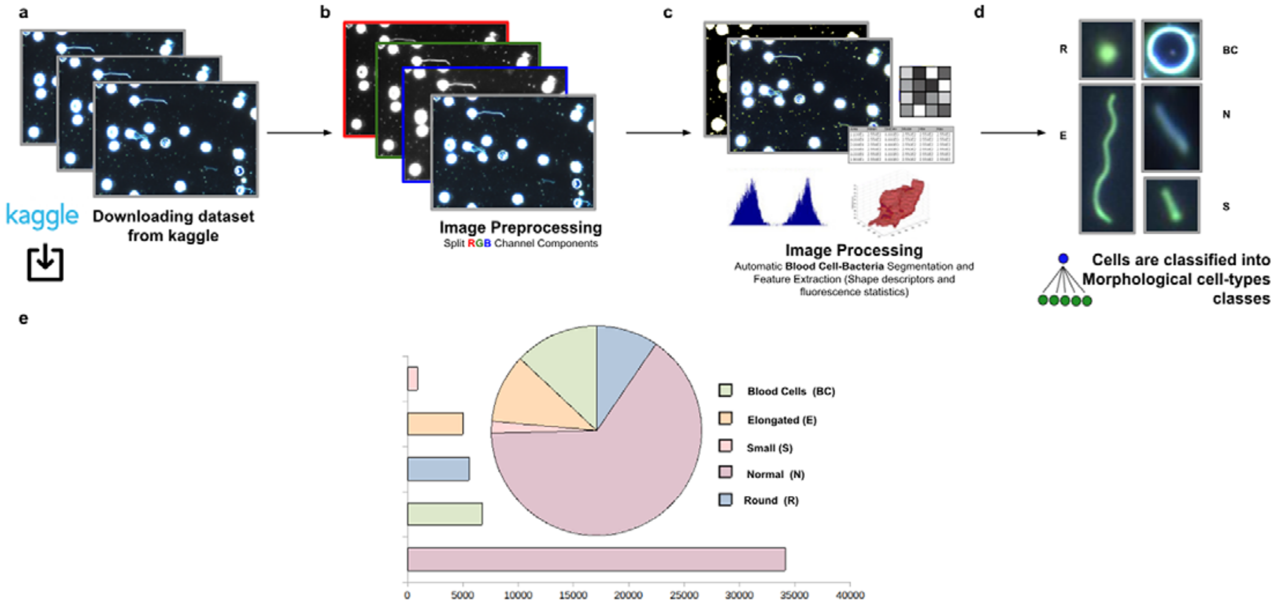

Figure 11. *Scheme of semiautomated analysis of raw images for classifying* Spirochaete *bacteria in the blood. (a) The full dataset of images to be analyzed was downloaded from Kaggle. (b) Image preprocessing actions to get the separated channel components were applied. (c) Image Processing actions for Blood Cells–Bacteria segmentation and subsequent identification. A vector describes each cell based on shape descriptors, geometry, and fluorescence statistics. (d) Cells were classified into Blood Cells (BC), Round (R), Elongated (E), Small (S), and Normal (N) cell-type classes. (e) Quantification results of classifying cells belonging to each morphological class.*

analysis pipelines[36]. In this sense, Cell-TypeAnalyzer introduces the batch processing to implement the cell-type analysis based on user-defined conditions on large image datasets (once the user is satisfied with single-image results), avoiding both the individual and tedious processing of each image. Regarding data visualization, such as with CPA software which offers heatmaps, boxplots, and histograms, Cell-TypeAnalyzer allows users for data visualization and exploration to easily drill down the cell-type classification results using dynamic scatter plots.

Regarding usability, the semiautomated analysis provided by Cell-TypeAnalyzer does not require any programming proficiency thanks to its user-friendly wizard-like GUI and its quite intuitive visualization settings. Nonetheless, some instructions and video tutorials are supplied in our documentation (https://github.com/QuantitativeImageAnalysisUnitCNB/CellTypeAnalyzer). As already happens with CPA software, Cell-TypeAnalyzer code is entirely open-source, and it does not require any commercial license. In terms of functionality, as in the case of CPA software, time-lapse data are not supported for analysis being solely possible to be used with static images; instead, Cell-TypeAnalyzer is able to process each slice from time-lapse image independently. CPA relies on CellProfiler to extract a huge amount of features to describe each cell presuming that user has prior knowledge of cell types present on images thus whether there is a large set of images, it is impossible to identify all the significant cell types by visual inspection. For these reasons, Cell-TypeAnalyzer computes features of each cell readily understandable for average users such us shape descriptors (perimeter, area, and roundness) or intensity-based statistics in each channel within each segmented cell compartment.

On the other hand, these days, biologists are increasingly becoming more qualified users, which is leading to a deeper understanding of the data. As opposed to machine learning approaches in which the user is requested to costly label many input cells to train the underlying classifier, Cell-TypeAnalyzer suggests cell-type classification based on simple rules on the calculated features. A further consideration of Cell-TypeAnalyzer related to CPA is the potential limitation of being restricted to maximum of three input image channels for cell-type analysis.

Overall, by leveraging a blend of semiautomatic and manual tools, Cell-TypeAnalyzer may achieve an accurate and efficient single-cell detection in images of not well-separated objects as well as high speed in batch processing, allowing identification of hundreds of cell types per minute. By contrast, in cases of images where objects are touching, segmentation may be a tough task; therefore, manual verification along with the adjustment of preprocessing actions is required. Finally, the batch processing will generate a folder for each processed image along with a summary folder per directory processed, which makes it possible to researchers easily examine outputs to determine how different are the identified cell types. The data tables will be saved using the common CSV file format, enabling its use in any spreadsheet application for further complex analyses, which makes the data accessible to researchers without programming experience. All together, these features make Cell-TypeAnalyzer, despite some limitations, an accessible plugin for researchers at different levels of domain which facilitates cell-type classification under user-defined conditions at different phases of the cell cycle (Figure 12).

5. Methods and Materials

5.1. Confocal and widefield microscopy

Images were acquired with a Confocal multispectral system Leica STELLARIS 5, using three laser lines: 405, 488, and 638 nm for DAPI, Alexa 488, and Alexa 647 excitation, respectively, and three Power HyD S spectral detectors for the fluorochromes' emission detection and integrated software module for real-time multidimensional super-resolution multidimensional image detection and processing (Lightning) and Leica DMi8 S widefield epifluorescence microscope with led lines: 405, 490, and 635 nm for DAPI,

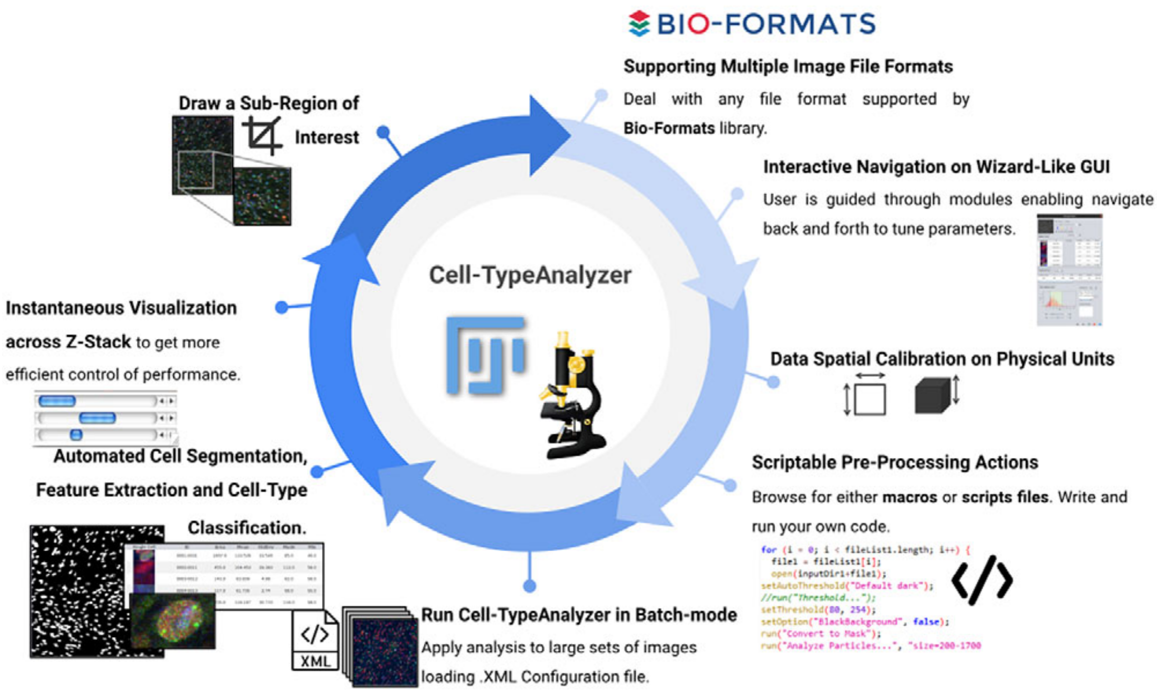

Figure 12. *Schematic description of Cell-TypeAnalyzer main functionalities. It is an open-source Fiji or ImageJ plugin for the semiautomated classification of cells according to specific cell types defined by the user. It offers a flexible and modular solution for users through an intuitive graphical user interface. It can deal with multiple image formats supported by the Bio-Formats library. It is also easily scriptable to perform preprocessing actions before cell segmentation and feature extraction. Cell-TypeAnalyzer allows the user to calibrate metrics on physical units, not in pixels, together with having instant visualization of each step of the analysis.*

https://doi.org/10.1017/S2633903X22000058 Published online by Cambridge University Press

Alexa 488, and Alexa 647 excitation with the appropriate filter cubes to detect the specific emission of the fluorochromes, and a Hamamatsu Flash 4 sCMOS digital camera for image detection.

5.2. Development and implementation

Cell-TypeAnalyzer was developed in Eclipse integrated development environment[37] for Java Developers version 2019-12 (4.14.0), an open-source platform mainly written in Java and used in computer programming for computer programming developing user-friendly Java applications. Cell-TypeAnalyzer is a Java application that inherits from ImageJ's plugin class, thus extending ImageJ's ecosystem. The core software and GUI were built using Java 8. Plots and histograms were implemented using the JFreeChart library. For reading the input images, we used the Bio-Formats library[19]. For handling XML files, we used JDom, and for handling Microsoft Office Formats (.xls and .xlsx), we used Apache POI libraries.

5.3. Installing in Fiji or ImageJ

Cell-TypeAnalyzer runs as a plugin of Fiji or ImageJ (https://imagej.nih.gov/ij/download.html) and consequently can be executed in Windows, Mac OS, or Linux systems. Cell-TypeAnalyzer plugin does not have an updated site yet. To install it, the file CellTypeAnalyzer_jar must be downloaded from https://github.com/QuantitativeImageAnalysisUnitCNB/CellTypeAnalyzer and moved into the ImageJ/Fiji plugins subfolder. Alternatively, it can be dragged and dropped into the ImageJ/Fiji main window or, optionally, installed through ImageJ/Fiji menu bar Plugins → Install → Path to File. After installing the plugin, ImageJ or Fiji must be restarted.

5.4. Supported image file formats

Cell-TypeAnalyzer deals with a wide range of file formats using Bio-Formats[19], an open-source library from life sciences supporting or reading almost any image format or multidimensional data as *z*-stacks, time series, or multiplexed images, keeping metadata easily accessible. In case of loading a Leica Image File[38] whose extension is .lif, which is a file format allowing storing several image series in the same file, our software is capable of extracting each image automatically as a single TIFF file, keeping the original pixel values and spatial calibration. On top of that, the user has access to a list of images that are available during the whole procedure for updating analysis as many times as needed. Regarding the limitation of usage, Cell-TypeAnalyzer is restricted to images in 2D single-plane RGB form: 24-bit RGB or Color Composite.

5.5. Code availability

Source code and documentation for the plugin are available at https://github.com/QuantitativeImageAnalysisUnitCNB/CellTypeAnalyzer.

6. Conclusions

This paper presents Cell-TypeAnalyzer, a plugin for automatically detecting and semiautomatically classifying cells according to very flexible cell-type definitions in multiple microscope image files. This tool was developed as a plugin working under both ImageJ and Fiji platforms. The implemented procedure consists of image preprocessing actions, cell segmentation, cell characterization through the extraction of features in RGB channels, and cell classification. This tool was designed to interactively guide users through various modules, allowing navigating back and forth to tune parameters or review processing actions while performing cell classification. Therefore, Cell-TypeAnalyzer offers a user-friendly, generic, and flexible strategy that can be applied to a wide range of biological challenges to examine relationships among cells that might reveal worthy new biological insights.

Acknowledgments. We are very grateful to members from the Advanced Light Microscopy facility at the National Center for Biotechnology, Ana Oña Blanco, and Gianluca D'Agostino for providing us the dataset of cells imaged by both the confocal and widefield microscopy and also for their collaborative help in this research. We are also thankful to Beatriz Serrano for pointing us to IDR and some of its datasets.

Competing Interests. The authors declare no competing interests exist.

Authorship Contributions. A.C.L., J.A.G.-P., and C.O.S.S. conceived the project and designed the algorithms. A.C.L. wrote the software code and performed all experiments. A.M.O.B. prepared the samples and acquired the images at the microscope. All authors wrote and revised the manuscript.

Funding Statement. This research was supported by the Spanish MICINN (PRE2018-086112) by the FPI fellowship from the Spanish Ministry of Science and Innovation through the Severo Ochoa excellence accreditation SEV-2017-0712-18-1. We would also like to acknowledge economical support from the Grant PID2019-104757RB-I00 funded by MCIN/AEI/10.13039/501100011033/ and "ERDF A way of making Europe," by the "European Union," the Grant SEV-2017-0712 funded by MCIN/AEI/10.13039/501100011033, and the European Union (EU) and Horizon 2020 through grant HighResCells (ERC—2018—SyG, Proposal: 810057).

Data Availability Statement. Source code and documentation for the plugin are available at https://github.com/QuantitativeImageAnalysisUnitCNB/CellTypeAnalyzer.

Supplementary Materials. To view supplementary material for this article, please visit http://doi.org/10.1017/S2633903X22000058.

References

1. Kervrann C, Sorzano CÓS, Acton ST, Olivo-Marin J-C & Unser M (2016) A guided tour of selected image processing and analysis methods for fluorescence and electron microscopy. *IEEE J Sel Top Signal Process* **10**, 6–30.
2. Lunde A & Glover JC (2020) A versatile toolbox for semi-automatic cell-by-cell object-based colocalization analysis. *Sci Rep* **10**, 19027.
3. Waters JC (2009) Accuracy and precision in quantitative fluorescence microscopy. *J Cell Biol* **185**(7), 1135–1148.
4. Schneider C, Rasband W & Eliceiri K (2012) NIH image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671–675.
5. Barry DJ, Durkin CH, Abella JV & Way M (2015) Open source software for quantification of cell migration, protrusions, and fluorescence intensities. *J Cell Biol* **209**(1), 163–180.
6. O'Brien J, Hayder H & Peng C (2016) Automated quantification and analysis of cell counting procedures using ImageJ plugins. *J Vis Exp* **117**, 54719.
7. Trapnell C (2015) Defining cell types and states with single-cell genomics. *Genome Res* **25**, 1491–1498.
8. Cossarizza A, Chan HD, Radbruch A, Acs A, Adam D & Adam-Klages S (2019) Guidelines for the use of flow cytometry and cell sorting in immunological studies. *Eur J Immunol* **49**(10), 1457–1973.
9. Vembadi A, Menachery A & Qasaimeh MA (2019) Cell cytometry: review and perspective on biotechnological advances. *Front Bioeng Biotechnol* **7**, 147.
10. Ferrand A, Schleicher KD, Ehrenfeuchter N, Heusermann W & Biehlmaier O (2019) Using the NoiSee workflow to measure signal-to-noise ratios of confocal microscopes. *Sci Rep* **9**, 1165.
11. Smal I, Loog M, Niessen WJ & Meijering E (2010) Quantitative comparison of spot detection methods in live-cell fluorescence microscopy imaging. *IEEE Trans Med Imaging* **29**(2), 282–301.
12. Schindelin J, Arganda-Carreras I, Frise E, *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**(7), 676–682.
13. Schindelin J, Rueden CT, Hiner MC & Eliceiri KW (2015) The ImageJ ecosystem: an open platform for biomedical image analysis. *Mol Reprod Dev* **82**(7–8), 518–529.
14. Alden AD, Elliott JT, Filliben JJ, *et al.* (2011) Comparison of segmentation algorithms for fluorescence microscopy images of cells. *Cytom Part A* **79**(7), 545–559.
15. Buggenthin F, Marr C, Schwarzfischer M, *et al.* (2013) An automatic method for robust and fast cell detection in bright field images from high-throughput microscopy. *BMC Bioinform* **14**, 297.
16. Zou T, Pan T, Taylor M & Stern H (2021) Recognition of overlapping elliptical objects in a binary image. *Pattern Anal Appl* **24**, 1193–1206.
17. Bankhead P, Loughrey MB, Fernández JA, *et al.* (2017) Qupath: open source software for digital pathology image analysis. *Sci Rep* **7**(1), 16878.
18. Bray MA, Singh S, Han H, *et al.* (2016) Cell painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nat Protoc* **11**, 1757–1774.
19. Linkert M, Rueden CT, Allan C, *et al.* (2010) Metadata matters: access to image data in the real world. *J Cell Biol* **189**(5), 777–782.

20. Meijering E (2012) Cell segmentation: 50 years down the road life sciences. *IEEE Signal Process Mag* **29**(5), 140–145.
21. Vicent L & Soille P (1991) Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Trans Pattern Anal Mach Intell* **13**(6), 583–598.
22. Legland D, Arganda-Carreras I & Andrey P (2016) MorphoLibJ: integrated library and plugins for mathematical morphology with ImageJ. *Bioinform* **32**(22), 3532–3534.
23. Lapytsko A, Kollarovic G, Ivanova L, Studencka M & Schaber J (2015) FoCo: a simple and robust quantification algorithm of nuclear foci. *BMC Bioinform* **16**, 392.
24. Ferreira T & Rasband W (n.d.). "ImageJ User Guide - IJ 1.46", imagej.nih.gov/ij/docs/guide/,2010-2012.
25. Tinevez J-Y, Perry N, Schindelin J, *et al.* (2017) Trackmate: an open and extensible platform for single-particle tracking. *Methods* **115**, 80–90.
26. Jonkman J & Brown CM (2015) Any way you slice it—a comparison of confocal microscopy techniques. *J Biomol Tech* **26**(2), 56–65.
27. Aknoun S, Yonnet M, Djabari Z, *et al.* 2021) Quantitative phase microscopy for non-invasive live cell population monitoring. *Sci Rep* **11**(1), 4409.
28. Tarnowski BI, Spinale FG & Nicholson JH (1991) DAPI as a useful stain for nuclear quantitation. *Biotech Histochem* **66**(6), 296–302.
29. Fragkos M, Jurvansuu J & Beard P (2009) H2AX is required for cell cycle arrest via the p53/p21 pathway. *Mol Cell Biol* **29** (10), 2828–2840.
30. Williams E, Moore J, Li SW, *et al.* (2017) Image data resource: a bioimage data integration and publication platform. *Nat Methods* **14**, 775–781.
31. Fuchs F, Pau G, Kranz D, *et al.* (2010) Clustering phenotype populations by genome-wide RNAi and multiparametric imaging. **8**(370), 775–781.
32. Pollehn L (2020) Bacteria detection with darkfield microscopy. Dataset for *spirochaete* segmentation with image and manually annotated mask, Version 1. https://www.kaggle.com/longnguyen2306/bacteria-detection-with-darkfield-microscopy/metadata (accessed June 2020).
33. Smith K, Piccinini F, Balassa T, *et al.* (2018) Phenotypic image analysis software tools for exploring and understanding big image data from cell-based assays. *Cell Syst* **6**(6), 636–653.
34. Carpenter AE, Jones TR, Lamprecht MR, *et al.* (2006) CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Cell Syst* **7**, R100.
35. Jones TR, Kang IH, Wheeler DB, *et al.* (2008) CellProfiler analyst: data exploration and analysis software for complex image-based screens. *BMC Bioinform* **9**, 482.
36. Dobson ETA, Cimini B, Klemm AH, Wählby C, Carpenter AE & Eliceiri KW (2021) ImageJ and CellProfiler: complements in open-source bioimage analysis. *Curr Protoc* **1**(5), 2424–2426.
37. desRivieres J & Wiegand J (2004) Eclipse: a platform for integrating development tools. *IBM Syst J* **43**(2), 371–383.
38. Besson S, Leigh R, Linkert M, *et al.* (2019) Bringing open data to whole slide imaging. In *Digital Pathology ECDP*, Lecture Notes in Computer Science, vol. 11435, pp. 3–10. Cham: Springer.
39. Long Pollehn, Bacteria detection with darkfield *microscopy* — Dataset for spirochaeta segmentation with *image* and manually annotated masks, 2020, Kaggle.com, https://www.kaggle.com/longnguyen2306/bacteria-detection-with-darkfield-microscopy, Data files © Original Authors

Cite this article: López A. C, Gómez-Pedrero J. A, Blanco A. M. O and Sorzano C. O. S (2022). Cell-TypeAnalyzer: A flexible Fiji/ImageJ plugin to classify cells according to user-defined criteria *Biological Imaging*, 2: e5. doi:https://doi.org/10.1017/S2633903X22000058

Appendix B

TrackAnalyzer: A Fiji/ImageJ Toolbox for a holistic Analysis of Tracks

Cayuela López, A., Gómez-Pedrero, J., Blanco, A., Sorzano, C. (2022). Cell-TypeAnalyzer: A flexible Fiji/ImageJ plugin to classify cells according to user-defined criteria. *Biological Imaging*, 2, E5. doi:10.1017/S2633903X22000058

Biological Imaging (2022), 1: 1–14
doi:10.1017/blg.2021.xx

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

TrackAnalyzer: A Fiji/ImageJ Toolbox for a holistic Analysis of Tracks

Ana Cayuela López[1]*, Eva M. García-Cuesta[2], Sofía R. Gardeta[2], José Miguel Rodríguez-Frade[2], Mario Mellado[2], José Antonio Gómez-Pedrero[3] and Carlos Oscar S. Sorzano[1]*

[1]Biocomputing Unit, National Centre for Biotechnology, Cantoblanco, Madrid, Spain
[2]Department of Immunology and Oncology, National Centre for Biotechnology, Cantoblanco, Madrid, Spain
[3]Applied Optics Complutense Group, Faculty of Optics and Optometry, University Complutense of Madrid, Madrid, Spain
*Corresponding author. Email: acayuela@cnb.csic.es; coss@cnb.csic.es

Received: 31 January 2021; **Revised:** 01 May 2021; **Accepted:** 06 May 2021

Keywords: spot; cell; track, single-particle tracking ; time-lapse; diffusion

Abstract

Current live-cell imaging techniques make possible the observation of live events and the acquisition of large datasets to characterise the different parameters of the visualized events. They provide new insights into the dynamics of biological processes with unprecedented spatial and temporal resolutions. Here we describe the implementation and application of a new tool called TrackAnalyzer, accessible from Fiji and ImageJ. Our tool allows running semi-automated Single-Particle Tracking (SPT) and subsequent motion classification, as well as quantitative analysis of diffusion and intensity for selected tracks relying on the GUI for large sets of temporal images (X-Y-T or X-Y-C-T dimensions). TrackAnalyzer also allows 3D visualization of the results as overlays of either spots, cells or end-tracks over time, along with corresponding feature extraction and further classification according to user criteria. Our analysis workflow automates the following steps: 1) spot or cell detection and filtering, 2) construction of tracks, 3) track classification and analysis (diffusion and chemotaxis), and 4) detailed analysis and visualization of all the outputs along the pipeline. All these analyses are automated and can be run in batch mode for a set of similar acquisitions.

Impact Statement

In recent decades, single-particle tracking analysis has become a powerful method to evaluate biomolecules' diffusion dynamics and interactions in living cellular ecosystems. Because changes in biomolecule dynamics can lead us to understand either functional states or signalling pathways, this tool allows characterizing the mechanisms of one molecule at a time within single trajectories by extracting mobility-related properties together with performing mean-squared displacement approaches to quantitatively analyze diffusion, thus getting further track classification. Here, we present TrackAnalyzer, a new open-source plugin which extends from TrackMate's single-particle tracking analysis broadly applicable under ImageJ or Fiji, which prevents users from using complex instruments and provides intuitive data analysis schemes hence leading users to a proper interpretation of information extracted from trajectories.

1. Introduction

With the development of breakthrough live-cell imaging techniques in optical microscopy, such as Confocal and Total Internal Reflection Microscopy (TIRF) over the last 40 years, quantitative analysis of dynamic processes at the sub-cellular level has become crucial to acquire valuable information related to dynamics intracellular processes over long periods of time with a spatial resolution of a few tens of nanometers[(1)]. In this context, due to advances in fluorescent protein labelling and software, single-particle tracking analysis (SPT) as a time-lapse imaging tool has become standard in life sciences to measure motion, diffusion properties, and the changing spatial distribution of single-particles in real-time with high-temporal resolution and high signal-to-noise ratio[(2)]. Since particles are fluorescently labelled, SPT analysis must seek to roughly reconstruct the motion of single particles of interest over consecutive time points. Its combination with markers allows monitoring vital cellular processes such as cell differentiation[3]. The estimation of frame-to-frame correspondence among particles at the cellular and molecular levels (with high accuracy and high reproducibility) requires a high signal-to-noise ratio. However, this is not always achieved due to severe noise from a fluctuating background, autofluorescence, blinking, photobleaching, phtototoxicity[4], poor contrast, extremely high particle density, or motion heterogeneity. To address these challenging events, inherent to the dynamic organization of cellular components (and essential for many biological processes) such as cell division, differentiation, cell adhesion, or migration[5], SPT algorithms explicitly take some countermeasures to guarantee the correct tracking of these fluorescent particles. Specifically: (I) gap-closing events when a single particle temporarily disappears from focus and reappears later; this type of event is closely related to fluorophore blinking and stochastic fluctuations of spot or cell intensity so the tracker algorithm may bridge missing detections in a predefined number of subsequent frames[6]; (II) merging events when two single particles approach each other and fuse into a unique object; (III) splitting events when a single-particle splits into two new single-particles[7]. To correctly compute trajectories, single-particle candidates considered as relevant must be accurately detected and isolated from each other and from a background with nanometer spatial and millisecond temporal resolution[8]. Thus, enhancing the signal-to-noise ratio is mandatory; the higher the background noise, the more distorted the tracking. SPT analysis involves spatial methods for (I) single-particle detection in which each spot or cell is segmented, identified, located and isolated from the background establishing X-Y-Z-T coordinate correspondences frame-by-frame, and temporal methods for (II) single-particle linking in which each single-particle detected is assigned over time into a single track.

While manual single-particle tracking of biological processes may be straightforward when the particle density is low, tracking large datasets of sparse living cells is often a subjective, barely reproducible, and tedious task. Consequently, its automation is very much appreciated. Due to experimental constraints, fully automated SPT approaches frequently perform poorly when the experimental conditions change. For this reason, the possibility of combining automation with user control[9] may facilitate the quantification of live cell events. At present, there is still a lack of user-friendly and comprehensive software for single-particle tracking to cope with the enormous amount of time-lapse microscopy acquisitions arising from quite different experimental conditions. Given the current situation, we decided to construct TrackAnalyzer to allow the user to set up sophisticated SPT analyses tailored to his/her experimental conditions, apply this analysis in batch mode to a large collection of similar acquisitions, and finally analyze the results. Our software is available through an open-source plugin for Fiji[10] or ImageJ[11]. TrackAnalyzer performs the detection of the spots or cells to follow, the construction of the tracks, quantitative diffusion analysis, trajectory analysis, cluster size analysis, and single-step photobleaching analysis (see Fig. 1). Our viewer allows 2D visualization of the spots (or cells) and tracks, spot/track filtering and classification into user-defined specific spot/track types.

For detecting and constructing the tracks, our software takes advantage of the previously published open-source software TrackMate[12], which is an extensible platform running for either Fiji or ImageJ, openly available and very well-documented. TrackMate provides algorithms for spot or cell detection, track construction (automated, semi-automated, and manual tracking), visualization, and subsequent feature extraction. In this way, TrackMate addresses both usability and flexibility to provide users with a user-friendly tool to tackle the complexity of this type of analysis.

For the classification of the different types of trajectories, we use TraJClassifer[13]. This software classifies trajectories into their respective motion types: normal diffusion (ND), anomalous diffusion (AD), confined diffusion (CD), and directed motion (DM). An interesting feature of TraJClassifier is that trajectories can be divided into segments, and the motion type of each segment can be analyzed.

TrackAnalyzer implements an algorithm to calculate the diffusion coefficients of each trajectory. The algorithm is based on the mean-square-displacement (MSD) curve as a function of the time lag of each trajectory (see Sec. 5 in Materials and Methods). The short-time lag diffusion coefficient (D_{1-4}) is also calculated by fitting the first used-defined points of the MSD curve[5]. MSD-based methods are reliable for short trajectories, but they may be error-prone in longer trajectories due to their non-linearity and lack of distinction between modes of motion[14]. To overcome this non-linearity and describe non-linear diffusion, the anomalous exponent or alpha value (α) is calculated by the power-law form of the MSD, indicating the nonlinear relationship of the MSD with time[15]. The exponent of this power function determines whether the motion is confined ($0<\alpha<0.6$), anomalous ($0.6<\alpha<0.9$), free ($0.9<\alpha<1.1$), or directed ($\alpha>1.1$)[5]. For long trajectories, the moment scaling spectrum (MSS) together with its slope (S_{MSS}) is introduced as a method to categorize the various modes of motion[5]. While MSD-based analysis uses only the second moment, which can mislead in judging the type of motion, MSS uses

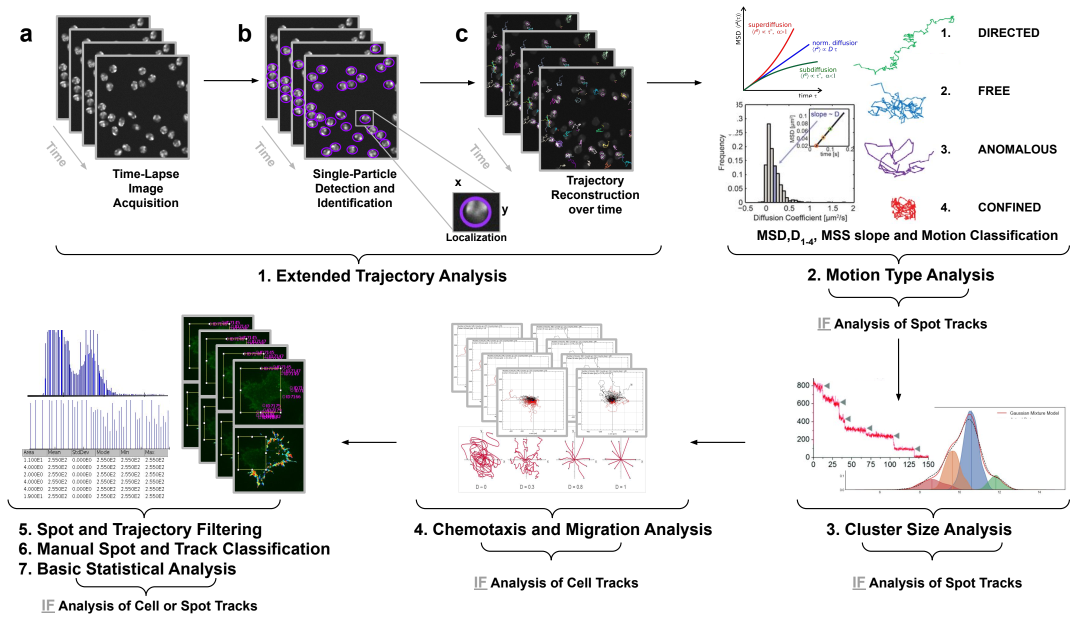

Figure 1. *Illustration of the workflow to perform single particle tracking together with subsequent analysis of diffusion using TrackAnalyzer software which consists of several processes.* ***(1)*** *Extended Trajectory Analysis.* ***a.*** *After Acquisition time series of multi-movie data sets.* ***b.*** *Localization, detection and subsequent identification of single particles framebyframe. A wide range of features is extracted based on the location, radius and image data.* ***c.*** *Single particles are linked to building trajectories over time (single particle tracking).* ***(2)*** *Motion Type Analysis. The resulting trajectories and links are analyzed after the tracking step to characterize them and evaluate the type of motion by applying quantitative analysis of diffusion, mean square displacement (MSD), and moment scaling spectrum (MSS) slope.* ***(3)*** *Cluster Size Analysis. The number of receptors per spot is calculated by applying Gaussian Mixture Model fitting and Single-step Photobleaching Analyses.* ***(4)*** *Chemotaxis and Migration Analysis. Several quantitative and statistical features (centre of mass, forward migration indices, velocity, ...) are calculated to characterize trajectories.* ***(5)*** *Spot and Trajectory Filtering,* ***(6)*** *Manual Spot and Track Classification, and* ***(7)*** *Basic Statistical Analysis. Features extracted from spots and tracks will be used to either filter or classify them depending on user-defined conditions..*

higher-order moments of the displacements. In this way, an S_{MSS} value of 0.5 defines Brownian or free motion, and S_{MSS} values below and above 0.5 determine confined and directed motion, respectively. Finally, a S_{MSS} of 0 determines immobility[(14)].

TrackAnalyzer also analyzes the spot or cell intensities along the whole trajectory. At this point, we provide the user with different algorithms to estimate the background fluorescence intensity described in Sec. 5.2 and use this estimated value to correct the raw measurement observed in the acquired images. This approach allows the identification of photobleaching. In combination with this approach, this tool provides an alternative strategy to evaluate the cluster size by fitting a Gaussian Mixture Model to the histogram of the logarithm of the background-subtracted integrated spot intensities.

Finally, we have also integrated the Chemotaxis and Migration Tool[(16)] to quantify both chemotaxis and migration experiments.

2. Results

2.1. Overview of the analysis procedure

The analysis workflow starts with the user calling TrackMate and setting up an analysis in this tool that correctly identifies the spots or cells and tracks in the specific experimental conditions of the dataset. TrackMate offers state-of-the-art segmentation and trajectory construction algorithms. After setting up the analysis, TrackMate will produce an XML file with the analysis parameters (this file also contains the results of the video analyzed, although these specific results are not of interest in our context of automated analysis of a collection of videos). The input to our software is the XML file produced by TrackMate, with the analysis parameters and the directory with the videos to be analyzed in batch mode. For each video in the input directory, we will create an output directory with the results of all

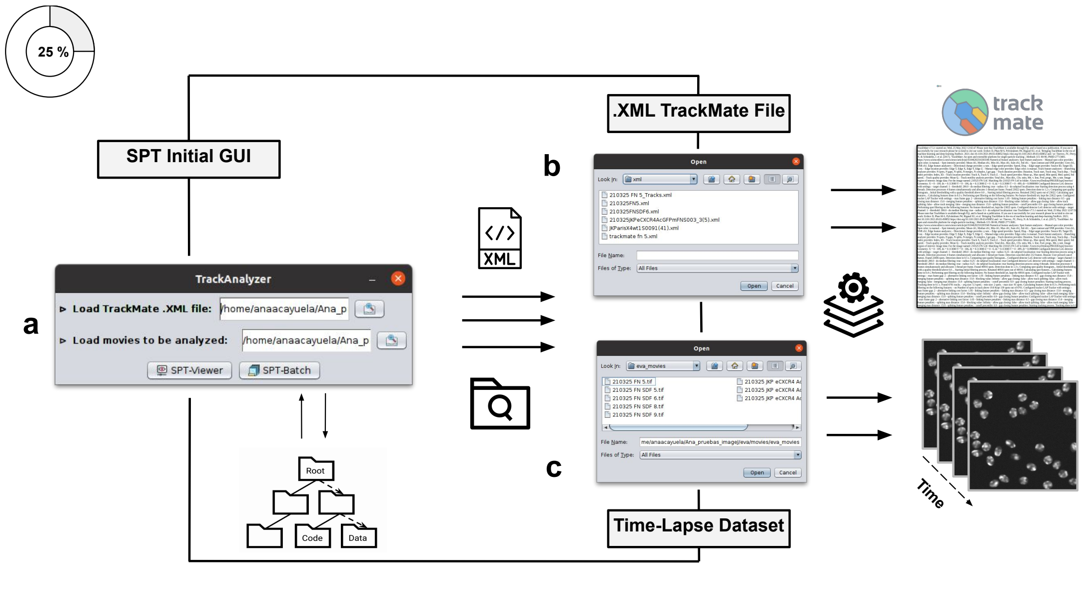

***Figure 2.** Illustration of getting started with the TrackAnalyzer plugin. **(a)** GUI structure of TrackAnalyzer. **(b)** TrackAnalyzer is started by selecting the .XML TrackMate configuration file and the time-lapse data sets to be analyzed in **(c)**..*

the different analyses on that video. Before launching the analysis in batch mode, the user must choose the parameters for all the different kinds of analyses performed on the tracks detected by TrackMate. Specifically:

1. **Extended trajectory analysis**. We provide a number of tools that help to extend the spot and trajectory analysis offered by TrackMate. In particular, the user may choose a specific frame range so that all spots detected out of this range are excluded. The user may also exclude all spots detected outside of a cell. We also offer different output options such as generating a summary file for each video, exporting the results in XML, text file or as a RoiSet that can be handled by ImageJ's RoiManager. Finally, the user may choose any XY scatter plot with information coming from the detected spots or cells, links or tracks-related features.

2. **Motion-type analysis**. TrackAnalyzer offers several ways of analyzing the motion type of the different trajectories. As a result, we classify trajectories into immobile or mobile depending on the threshold set by the user, and within these into confined, anomalous, free or Brownian, or directed trajectories. We calculate the short-time lag diffusion coefficient (D_{1-4}), Mean Squared Displacement (MSD) and diffusion coefficient for all trajectories (the formal definition of all these quantities are given in Sec. 5.1). These measurements are especially well-suited to short trajectories and characterise the movement's onset. Additionally, trajectories are classified into short and long trajectories, depending on a threshold given by the user. Long trajectories are further analyzed using the Moment Scaling Spectrum (MSS), better suited to account for their non-linearities. Finally, we have also integrated TraJClassifier that allows the local classification of the trajectory motion type, i.e., a spot may behave in one way during the first half of the trajectory and in another way in the second half. TraJClassifier's classification is based on a random forest trained on simulated data with different kinds of motion.

3. **Cluster size analysis.** This analysis tries to estimate the number of fluorophores at each spot. This information is very useful for identifying the presence of monomers, dimers, trimers, etc., within a cluster. An unbiased estimate should account for the background fluorescence, which must be subtracted before further analysis. TrackAnalyzer offers several methods to estimate the background, described in Sec. 5.2. In particular, we use the following two methods to estimate the number of fluorophores:

   (a) a Gaussian Mixture Model fit of the histogram of the logarithm of the background-subtracted integrated spot intensities.

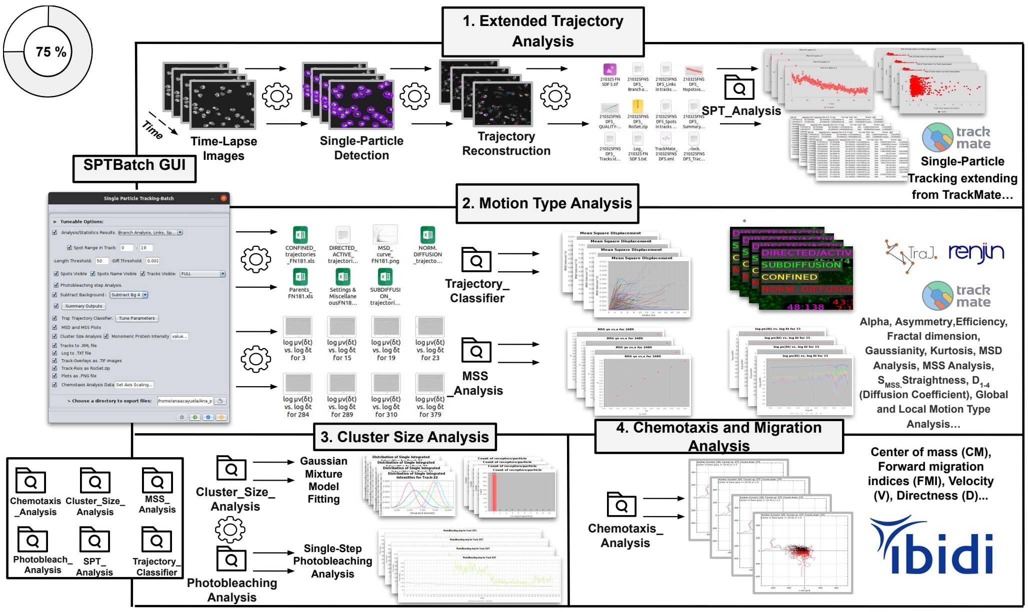

Figure 3. *Schematic overview of the SPTBatch procedure for single-particle tracking along with subsequent motion trajectory analysis, cluster size and single-step photobleaching analysis together with chemotaxis analysis in batch mode.* ***(1)*** *Extended Trayectory Analysis. Single-particle tracking analysis extending from TrackMate running in batch mode using multiple sets of files.* ***(2)*** *Motion Type Analysis. Trajectory analysis is executed to calculate short-time lag diffusion coefficient, diffusion coefficient, mean squared displacement curve, motion type classification, ...* ***(3)*** *Cluster Size Analysis and Single-step Photobleaching analysis is run.* ***(4)*** *Chemotaxis and Migration Analysis to quantify chemotactic cell migration..*

(b) a single-step photobleaching analysis. This technique analyzes the time evolution of the fluorescence of an individual spot along its trajectory. The number of photobleaching steps over time is a lower bound of the number of fluorophores in the spot.

4. **Chemotaxis and migration analysis.** The identified trajectories can be subjected to a chemotactic and migration analysis (as implemented by the Chemotaxis and Migration tool of ImageJ[(16,17)]). This tool allows the quantitative and statistical analysis of the migration of the spot centre of mass (CM) and the calculation of the forward migration indices (FMI), velocity (V), and directness (D) (described in 7, 6 and 8 in Sec. 5.1).

5. **Spot and trajectory filtering.** The user may explore the results once the batch analysis has been performed on all videos. This is done through a visualization tool that allows navigating the spots and tracks, showing information about their location in space and time and quality measures (as reported by TrackMate). This information is displayed as an interactive table. Clicking on any of its rows automatically brings us to the selected spot and track within the video.
   At this point, the user may further filter the results by applying specific criteria:

   - Spots and tracks can be manually selected/deselected for further statistical analysis.
   - The user can manually draw a region in the video and select/deselect spots and trajectories inside or outside that region.

6. **Manual trajectory classification.** Additionally, spots and tracks can be categorized by visually setting thresholds on the histogram of any of the features displayed in the table. Categories or classes can be defined by an arbitrary number of features (see Fig. 4) to determine specific spot and track types for further classification.
   The table of selected spots and tracks, along with their characterization in terms of spatial and time location and different descriptors, can be exported as a CSV file for further analysis outside our tool.

7. **Basic statistical analysis.** The final step of our tool offers basic exploratory statistical operations. For instance, the user may construct XY plots with any features calculated for the spots and tracks. These plots can be done only for one specific trajectory category (see the previous point in the workflow) or for all of them with their category used as a colour. Histograms of the different features can also be calculated, and basic statistical descriptors (mean, standard deviation, minimum, maximum, quantiles, ...) are given.

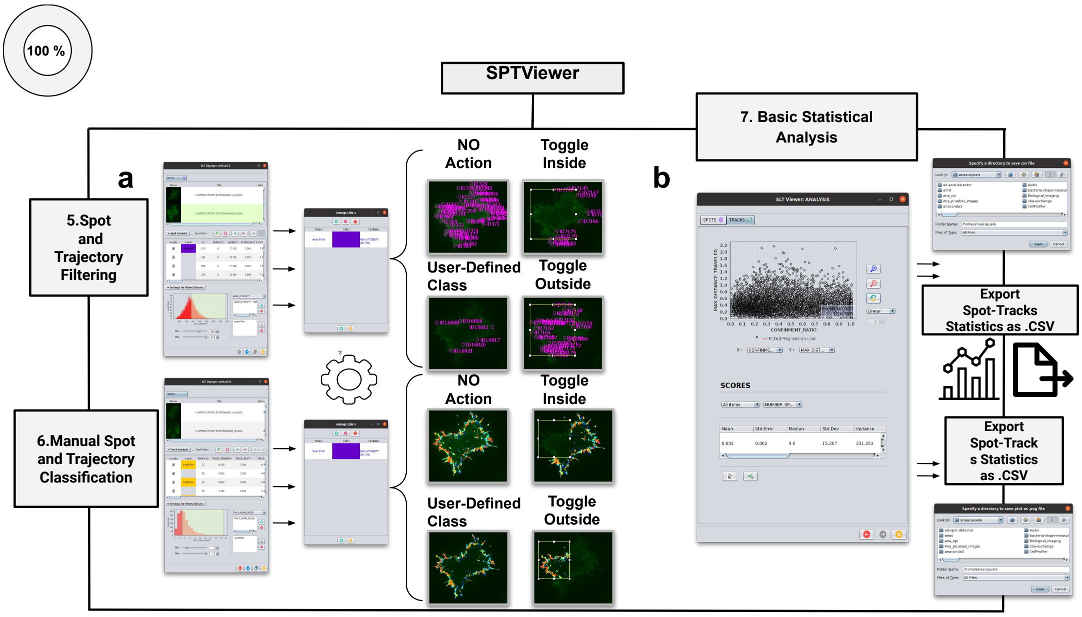

Figure 4. *Schematic overview of the manual analysis for spot and trajectory filtering.**(a)** The double tabbed wizard-like GUI of our viewer in which the user can configure the settings for either spot or trajectory filtering along with user-definition of classes to identify specific spot or track types retaining. **(b)** SPTViewer last wizard enables to configure dynamic scatter plots to display any spot/track feature as a function of any other.*

2.2. *Validation of the Method*

2.2.1. *Experimental dataset 1: Analysis of spot tracks*

Analysis of the dynamic of CXCR4 at the plasma membrane of Jurkat CXCR4$^{-/-}$ cells electroporated with CXCR4-AcGFPm. In this example, we will illustrate the features that TrackAnalyzer offers for the different kinds of analysis of the tracks of spots. In particular, Steps 1, 2 and 3 (see Sec. 2.1 in Overview of the analysis procedure). Two datasets with Jurkat CXCR4$^{-/-}$ cells electroporated with CXCR4-AcGFPm were used. This cell line, derived from human T lymphocytes, was generated using the CRISPR/Cas9 system to eliminate the endogenous expression of CXCR4[(18)]. Therefore, these cells only express CXCR4 labelled with AcGFPm. Cell sorting allowed us to obtain cells expressing 8,500 to 22,000 receptors per cell to work in single-particle conditions. It has been reported that ligand CXCL12 stimulation promotes CXCR4 nanoclustering at the cell membrane, which is necessary for a correct cell function[(19)]. In a previous work[(19)], particles in TIRF images were automatically detected, tracked and analyzed using described algorithms for diffusion analysis[(20)] implemented in MATLAB. We now compare the results obtained in our previous work with TrackAnalyzer's results.

The image sets studied in this case consist of time-lapse images acquired by a TIRF microscope (Leica AM TIRF inverted; 100x oil-immersion objective HCX PL APO 100x/1.46 NA) equipped with an EM-CCD camera (Andor DU 885-CS0-10-VP), at 37 °C with 5% CO_2. Image sequences of individual particles (500 frames) were then acquired at 49% laser (488-nm diode laser) power with a frame rate of 10 Hz (100 ms per frame). The penetration depth of the evanescent field used was 90 nm. The first dataset contains image sequences from 18 different cells on fibronectin (basal) conditions (Fig. 5a), while the second dataset contains image sequences from 14 different cells on fibronectin+CXCL12 (stimulated) conditions (Fig. 5b).

Before entering into TrackAnalyzer, we generated an XML parameter file with TrackMate. Spots were identified through subpixel localization applying LoG (Laplacian of Gaussian) detector[(12)]

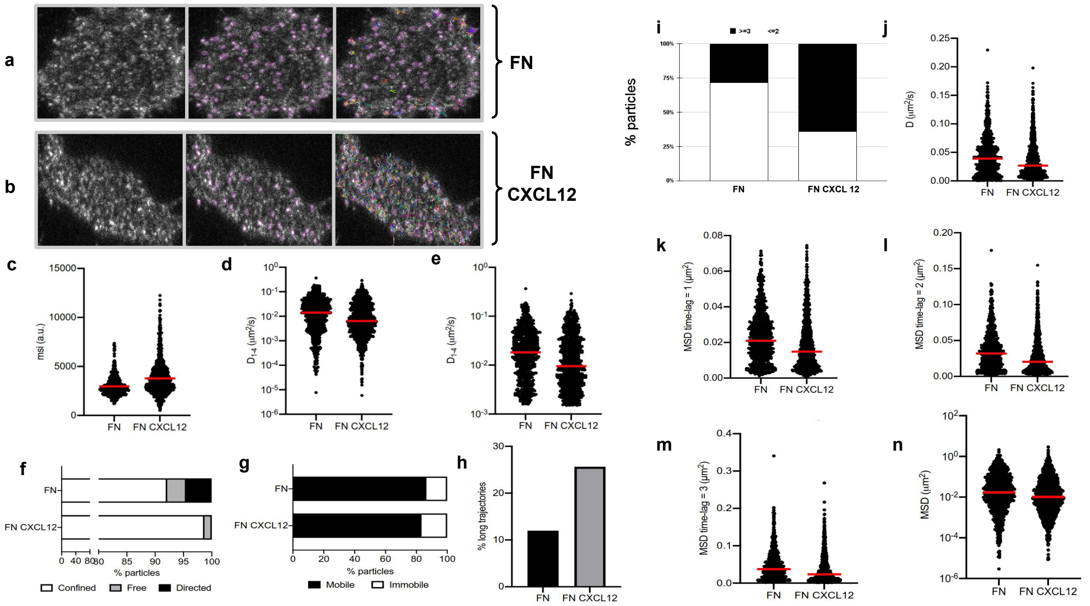

Figure 5. *Application of TrackAnalyzer to track CXCR4-AcGFPm in JK CXCR4$^{-/-}$ cells electroporated with CXCR4-AcGFPm. **(a-b)** Images of Jurkat CXCR4$^{-/-}$ cells electroporated with CXCR4-AcGFPm on fibronectin (FN)- **(a)** and FN+CXCL12-coated coverslips **(b)**. Scale bar, 5 μm. **(c-f)** Tracking results from TrackAnalyzer (741 particles in 18 cells on FN and 1,209 particles in 14 cells on FN+CXCL12). **(c)** Mean Spot Intensity (MSI, arbitrary units, a.u.) from individual CXCR4-AcGFPm trajectories. The mean is indicated (red). Short-time lag diffusion coefficients (D_{1-4}) of all **(d)** and mobile **(e)** single trajectories. The median is indicated (red). (*** p≤0.001, **** p≤0.0001, Welch's t-test).**(f)** Percentage of confined, free and directed CXCR4-AcGFPm particles at the cell membrane using the slope of MSS. **(g)**Percentage of mobile and immobile CXCR4-AcGFPm particles at the cell membrane. **(h)** Percentage of long trajectories of CXCR4-AcGFPm particles at the cell membrane. **(i)** Frequency of CXCR4-AcGFP particles containing the same number of receptors [monomers plus dimers (≤2) or nanoclusters (≥3) in cells, calculated from MSI values of each particle as compared with the MSI value of monomeric CD86-AcGFP. **(j)** Diffusion coefficients (D) of single trajectories. The median is indicated (red). **(k)** Mean Squared Displacement (MSD) of single trajectories using the first-time lag. The median is indicated (red). **(l)** Mean Squared Displacement (MSD) of single trajectories using the second time lag. The median is indicated (red). **(m)** Mean Squared Displacement (MSD) of single trajectories using third time-lag. The median is indicated (red). **(n)** Mean Squared Displacement (MSD) of single trajectories using more than three time-lags. The median is indicated (red)..*

(estimated object diameter=0.5 μm, quality threshold=500, Sub-pixel localization=true, Median filtering=true). Frame-to-frame spot linking was performed using TrackMate's LAP (Linear Assignment Problem) by closing gaps (linking max distance=0.5 μm; track segment gap closing=0.1 μm and 6 frames; track filtering of those trajectories of at least 20 frames).

We then launched TrackAnalyzer in batch mode to analyze all videos in the datasets with the same parameters. The following paragraphs provide the parameters and describe the results of the different kinds of analyses.

1. **Extended trajectory analysis**. We did not discard any of the identified tracks. As can be seen in Fig. 5c, stimulation with CXCL12 promotes an increase in the mean spot intensities (MSI) mean value of CXCR4 particles (2,970 arbitrary units for fibronectin vs. 3,781 arbitrary units for fibronectin+CXCL12) reflecting an increase of larger CXCR4 nanoclusters.
2. **Motion-type analysis**. A diffusion coefficient of 0.0015 was set as the threshold to discriminate among mobile and immobile particles, which is the percentile 95 of the diffusion coefficients of purified AcGFPm protein particles immobilized on glass coverslips[19]. Figs. 5d, e, g, j, k, l, m and n show the D_{1-4}, D, MSD, and percentage of the immobile particles. There is an increase in the percentage of immobile particles in CXCL12-stimulated conditions (6,12% for fibronectin vs 10,40% for fibronectin+CXCL12). Mobile particles also showed a reduction in the MSD

and D_{1-4} demonstrating a significant reduction in overall receptor diffusivity (0,012 $\mu m^2/s$ for fibronectin vs 0,007 $\mu m^2/s$ for fibronectin+CXCL12). These results are consistent with those previously obtained using Matlab[(18,19)], in which CXCL12 stimulation promoted the formation of larger nanoclusters of CXCR4 that also showed a different dynamic behaviour as compared with the receptor in basal conditions.

We classified the trajectories whose length is larger than 50 frames into confined, anomalous, Brownian (free) or directed (Fig. 5f)using the moment scaling spectrum (MSS) described in Sec. 5.1 along with the percentage of long trajectories per condition.

3. **Cluster size analysis**. We used the background subtraction method 4 (described in Sec. 5.2). The total number of receptors per particle was assessed by dividing the mean particle intensity by the particle arising from monomeric protein, i.e. CD86-AcGFP, estimated through the analysis of spots with just one photobleaching step. Therefore, this value was used as the monomer reference to estimate the number of receptors or molecules per particle, as shown in Fig. 5i.

2.2.2. Experimental dataset 2: Analysis of cell tracks

Analysis of the directed cell migration capacity of Jurkat cells. In this section, we illustrate the chemotaxis and migration analysis module. To do so, we only use Steps 1 and 4 (see Sec. 2.1 in Overview of the analysis procedure). Note that some of the steps only apply to spots and not cells, for instance, Steps 2 and 3.

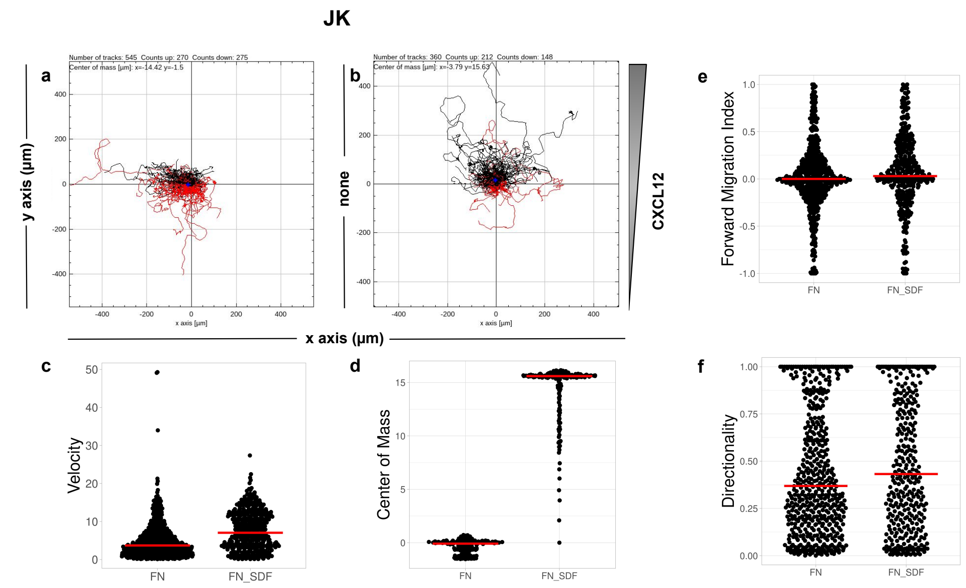

Figure 6. *Migration of JK cells in response to a CXCL12 gradient.* ***(a-b)*** *Representative spider plots showing the trajectories of tracked cells migrating along the gradient (black) or moving in the opposite direction (red). Black and red dots in the plots represent the final position of each single-tracked cell. The grey triangle indicates CXCL12 gradient. Quantification of the velocity* ***(c)****, center of mass* ***(d)****, forward migration index* ***(e)*** *and directionality* ***(f)*** *of experiments performed.*

4. **Chemotaxis and migration analysis**. In this analysis, we will illustrate other features of TrackAnalyzer to evaluate directional cell migration. Two datasets with Jurkat cells were used. To assess the ability of these cells, which express CXCR4 endogenously, to migrate toward CXCL12 gradients, we used fibronectin-coated chemotaxis chambers (Ibidi μ Slide Chemotaxis System; 80326). As CXCL12 is the ligand of CXCR4, we expected that cells migrate toward the gradient. The image sets studied in this case consist of time-lapse images acquired by Microfluor inverted microscope (Leica) every 2 minutes for 6 h at 37°C with 5% CO2. Single-cell tracking analysis was performed using TrackMate to generate an XML parameter file. Cells were identified through subpixel localization applying LoG (Laplacian of Gaussian) detector (estimated object diameter=0.5 μm, quality threshold=15.0, Sub-pixel localization=true, Median filtering=true).

Frame-to-frame cell linking was performed using TrackMate's LAP (Linear Assignment Problem) by closing gaps (linking max distance=60 μm ; track segment gap closing=60 μm and 2 frames). Then we launched TrackAnalyzer to analyze all videos in the datasets with these parameters. Therefore, directional cell migration was assessed (Fig. 6) by evaluating the corresponding spider plots (representing the trajectories of the tracked cells) (Fig. 6 a,b), forward migration index (FMI) (Fig. 6 e), directionality (D) (Fig. 6 f), the centre of mass (CM) (Fig. 6 d) and velocity (V) (Fig. 6 c) provided by chemotaxis and migration plugin integration. Quantitation of the results showed that JK cells sensed the CXCL12 gradient to increase the forward migration index, directionality, centre of mass and velocity.

Note that the steps 5, 6 and 7 are not presented in this section because its applicability does not address the biological questions arising from this context. Rather, they are deeply described in detail in Sec. 2.1 in Overview of the analysis procedure.

3. Discussion

TrackAnalyzer extends the existing tools for Single Particle Tracking analysis in two ways:

1. Batch-mode analysis. Most existing tools in ImageJ and Fiji allow the analysis of a single time-lapse dataset. However, many users and facilities do not have a single dataset but many datasets to analyze. Our tool allows the automatic analysis of all of them by configuring the analysis in one of them and replicating the same analysis to all other videos within the same experiment relying on the GUI.
2. Extended analysis. Most existing tools in ImageJ and Fiji specialize in a particular aspect of the tracks, for instance: TrackMate in identifying the spots and linking them into tracks; TraJClassifier in identifying their motion; Chemotaxis and Migration Tool in analyzing their motion from a different perspective. However, the user is also interested in other features like cluster size, measuring the motion in multiple ways, classifying the tracks into different categories and comparing their different features as a function of their categories, and removing from the analysis those tracks that have been incorrectly identified or focusing the analysis in a particular region of the cell. Our tool builds upon existing powerful tools and adds newly implemented measures to allow a more thorough analysis of all the tracks recorded in an experiment. In this way, we allow a very rich analysis of the particles' behaviour under various experimental conditions and allow a quantitative comparison of the different parameters that characterize the particles.

A track analysis's strength is correctly identifying the spots and their linkage to tracks. This is a rather challenging task that, if performed incorrectly, totally ruins the automatic analysis. TrackMate is extremely flexible in this aspect. It provides many different algorithms for spot identification, all of them fully configurable through a myriad of parameters (although the default values of most of them already give good results). TrackMate is also very strong and flexible in constructing the tracks from the set of spots. It also offers several highly configurable algorithms. In this regard, we consider that a semiautomated approach in which the user makes sure to configure the spot and track detection for his/her experimental conditions is crucial. This step is the key to the success of all the subsequent analyses. We have decided to rely on TrackMate for this identification step, as it is one of the most successful and adaptable programs available.

Icy[21] could have been an alternative to TrackMate. Icy is a free and open-source software for image analysis mainly oriented toward analysing biological images with a modular design composed of a kernel and plugins. Icy software integrates the Spot Tracking plugin[22], which ships automated methods for extracting tracks (particle tracking) from multiple objects (particle detection) as well as Track Manager plugin which provides relevant information from them (track analysis) in a sequence of 2D or 3D images. Track Manager allows the use of DSP-like trackProcessors enabling the display of tracks, time-based or ROI-based selection, and the generation of various views such as overlaid and animated local flow and polar graphs. These tools afford track filtering, classification (split tracks into tracklets to further statistically classify as Brownian/confined or directed), characterization by extracting features (confinement ratio, displacement distance, life time, intensity profile, instant speed, MSD, interaction analysis among tracked objets...) together with post-processing (export tracks into CSV files). This is a powerful tool to accurately perform common single-particle tracking analyses but compared with the integration of TraJClassifier, Icy may lack advanced track analysis capabilities. TraJClassifer provides diffusion characterization thought TraJ library and subsequent track classification by using simulated tracks of normal diffusion, subdiffusion, confined diffusion and directed motion. Then a group of features is estimated for each track, which together with the corresponding class, are used to train a random forest approach by means of Renjin. This extended track analysis also supports local analysis by splitting track into single segments with different motion types.

TrackAnalyzer benefits from the ImageJ ecosystem, probably the most known, flexible and longest-lived software for biomedical sciences. Consequently, TrackAnalyzer leverages from a lot of plugins for scientific image processing (as we have already done by integrating TrackMate with TraJClassifier and the Chemotaxis and Migration tool). To the best of our knowledge, TrackAnalyzer is the first tracking program within ImageJ that enables users to characterize and classify trajectories by a large number of descriptors, including the intensity and length of the tracks, multiple characterizations of their motion, cluster size by various methods, and their chemotactic features. Some protocols to quantitatively assess the tracks' motion, cluster size and intensity analysis were already designed in our previous work[20]. However, TrackAnalyzer now largely supersedes our analysis capacity.

4. Conclusions

In this paper, we have introduced TrackAnalyzer, a new Java-based plugin, an open-source and user-friendly toolkit to perform SPT analysis of multidimensional data in batch mode. This plugin operates equally well under ImageJ or Fiji ecosystems extending from TrackMate algorithms for (I) spot detection and spot analysis in which each spot receives a wide range of features based on its location, radius and metadata information; (II) linking spots together to build trajectories and get the subsequent trajectory analysis; (III) post-processing actions after SPT analysis such as 2D visualization and user-defined filtering of spots and trajectories. Our approach is semiautomatic as the user needs to define the TrackMate workflow to identify the spots. This strategy makes us capable of dealing with challenging experimental scenarios such as low signal-to-noise ratios or strong fluorescence backgrounds. In addition to the standard track analysis offered by TrackMate, we have included multiple ways of filtering the detected spots and tracks and various characterizations of their motion type, cluster size, chemotaxis and migration properties.

5. Materials and Methods

5.1. Motion analysis

Calculation of Mean Squared Displacement (MSD)

The MSD is the most common approach for analysing single-particle tracks[5]. Let us call Δt to the time difference between one frame in the time-lapse video and the next. The MSD of the particle j with time lag $n\Delta t$ is defined as[13]:

$$MSD_j(n\Delta t) = \frac{1}{N_j - n} \sum_{n'=1}^{N_j - n} \|\mathbf{r}_j((n' + n)\Delta t) - \mathbf{r}_j(n'\Delta t)\|^2 \tag{1}$$

where $\mathbf{r}_j(n'\Delta t)$ is the 2D location of the j-th particle at time $n'\Delta t$, and N_j is the length of the j-th trajectory in frames.

Calculation of Diffusion Coefficient (D)

The diffusion coefficient (D) is defined as the slope of the linear fitting of the first time lag of the MSD curve:

$$MSD(n\Delta t) = \Delta_0 + 4Dn \quad n = 1 \tag{2}$$

Calculation of the Short-Time Lag Diffusion Coefficient (D_{1-N})

The short-time lag diffusion coefficients (D_{1-N}) are defined as the slope of the linear fitting of the first N time lags (defined by the user) of the MSD curve:

$$MSD(n\Delta t) = \Delta_0 + 4D_{1-N}n \quad n = 1, 2, ...N - 1 \tag{3}$$

Calculation of the Anomalous Exponent (α)

The mean-square displacement plots must be fitted according to the anomalous diffusion model by power law fitting[5] according to:

$$MSD(n\Delta t) = \Delta_0 + 4Dn^{\alpha} \tag{4}$$

being α the anomalous exponent. The particle motion-type is classified as confined ($0 < \alpha < 0.6$), Brownian or free ($0.9 < \alpha < 1.1$) or directed ($\alpha > 1.1$)[5].

5.1.1. Calculation of Moment Scaling Spectrum (MSS) and its slope, (S_{MSS})

In the case of long trajectories, the moment scaling spectrum (MSS)[15,23,24] and its slope (S_{MSS}) was proposed as an approach to improve the calculation of MSD for non-linear diffusion. For each trajectory j the moments of displacement ($\mu_{j,\nu}$) were calculated for $\nu = 1, ..., 6$ as a function of time according to:

$$\mu_{j,\nu}(n\Delta t) = \frac{1}{N_j - n} \sum_{n'=0}^{N_j-n-1} \|\mathbf{r}_j((n'+n)\Delta t) - \mathbf{r}_j(n'\Delta t)\|^{\nu} \tag{5}$$

where r_j designates the position vector of track j at the time $n\Delta t$ being Δt the time interval and n the frame number $n = 0, 1, ...N_j - 1$ being N_j the trajectory length. The MSD is just a special case with $\nu =$ 2. In our implementation, we calculate all moments from $\nu = 1$ to $\nu = 6$ for each trajectory by plotting ($\mu_{j,\nu}$) against $n\Delta t$ in a double logarithmic plot, getting the scaling moments $\gamma_{j,\nu}$ from assuming each moment μ depends on the time shift according to $\mu_\nu(n\Delta) \sim n\Delta t^{\gamma\mu}$ [14,15]. Therefore plotting γ_ν against ν gives the moment scaling spectrum (MSS) and its slope (S_{MSS}) from linear regression characterizes the type of motion[25]: free ($S_{MSS} = 0.5$), directed (S_{MSS}>0.5), immobile ($S_{MSS} < 0.5$).

5.1.2. Calculation of Forward Migration Index ($FMI^{\parallel}, FMI^{\perp}$)

The Forward Migration Index (FMI) is an important measure for directed, chemotactic cell migration. It represents the efficiency of the forward migration of cells in the direction of a chemical gradient, **u**. We assume that **u** has unit length, and we also consider a direction perpendicular to **u** that will be referred to as $\mathbf{u}^{\perp}$. For a given particle, j, let $\mathbf{r}_j(0)$ and $\mathbf{r}_j(N_j\Delta t)$ be the first and last locations of its trajectory. The efficiency of the displacement in both directions are

$$\begin{aligned} FMI_j^{\parallel} &= \frac{\langle \mathbf{r}_j(N_j\Delta t) - \mathbf{r}_j(0), \mathbf{u} \rangle}{d_j} \\ FMI_j^{\perp} &= \frac{\langle \mathbf{r}_j(N_j\Delta t) - \mathbf{r}_j(0), \mathbf{u}^{\perp} \rangle}{d_j} \end{aligned} \tag{6}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product, and d_j is the total length of the j-th trajectory. The FMIs must be between -1 and 1. The larger the FMI in absolute value, the stronger the chemotactic effect is on the direction being studied. Finally, for a whole video, the FMI in a particular direction, parallel or perpendicular, is defined as the average of the corresponding particle FMIs.

5.1.3. End Center of Mass ($\mathbf{r}_{end}$)

The centre of mass represents the average of all single-cell endpoints. Its x and y values indicate the direction in which the group of cells primarily travelled.

$$\mathbf{r}_{end} = \frac{1}{J} \sum_{j=1}^{J} \mathbf{r}_j(N_j\Delta t) \tag{7}$$

where J is the total number of cells and $\mathbf{r}_j(N_j\Delta t)$ are the coordinates of the endpoint of each cell.

5.1.4. Directness (D)

The directness or directionality measures the straightness of the cell trajectories. For each cell, it is calculated by comparing the Euclidean distance and the accumulated distance between the starting point and the endpoint of a migrating cell:

$$D_j = \frac{\|\mathbf{r}_j(N_j\Delta t) - \mathbf{r}_j(0)\|}{d_j} \tag{8}$$

The directness values are always positive. A directness of $D = 1$ equals a straight-line migration from the start to the endpoint. The directness of a video is the average of the directness of all its cells.

5.2. Estimation of the background fluorescence

We now describe the different methods that we propose to estimate the background fluorescence.

- **Subtract Bg 1 (Manual).** Manual identification for each frame. This method enables user to manually select an indefinite number of positions over the Z-Projection image to ensure that the mean intensity measured belongs exclusively to areas within cell outside spots. This approach measures the mean background intensity for each frame for all selected locations along the video.
- **Subtract Bg 2 (Spot Ring).** This approach estimates the mean background intensity of each spot. It measures the intensity in a ring ranging from its radius to twice its radius.
- **Subtract Bg 3 (Inside the cell, not excluding spots).** This approach measures the mean background intensity for each frame by identifying the cell in each frame based. Then, the background mean intensity value is computed as the average within these masks (not excluding spots).
- **Subtract Bg 4 (Inside the cell excluding spots).** This approach measures the mean background intensity for each frame by identifying the cell in each frame based. The background is estimated as the average intensity within the cell, excluding the spot positions.
- **Subtract Bg 5 (Rolling ball).** This method estimates a locally varying background as the average within a rolling ball(26). It is important to note that the ball radius must be larger than the radius of the largest spot in the image.

5.3. Development and Implementation

TrackAnalyzer was developed in the Eclipse Integrated Development Environment (IDE)(27) for Java Developers version 2019-12 (4.14.0), an open-source platform mainly written in Java and used in computer programming for computer programming developing user-friendly Java applications. Each plugin is a Java application that inherits from ImageJ's plugin class extending from the TrackMate ecosystem. The core software and graphical user interface were built using Java 8. Plots and histograms were implemented using the JFreeChart library. For reading the input images, we used the Bio-formats library(28). For handling XML files, we used JDom, and for taking Microsoft Office Formats (.xls,.xlsx), we used Apache POI libraries. In the case of classifying trajectories, we called the TraJ Java library for diffusion trajectory (2D) analysis.

The source code and documentation are available at https://github.com/QuantitativeImageAnalysisUnitCNB/TrackAnalyzer_.

5.4. Installation in Fiji or ImageJ.

TrackAnalyzer must be installed as a plugin of Fiji or ImageJ (https://imagej.nih.gov/ij/download.html) and consequently can be executed in Windows, Mac OS, or Linux systems. The next step is to install TrackAnalyzer, which can be done by download the plugin from http://sites.imagej.net/TrackAnalyzer/plugins/ and moved into the ImageJ/Fiji `plugins` subfolder. Alternatively, it can be dragged and dropped into the ImageJ/Fiji main window or installed through ImageJ/Fiji menu bar `Plugins → Install → Path to File`. After installing the plugin, ImageJ or Fiji must be restarted. Note that to visualize the wizard-like GUI that guides the user through the set of predefined steps in this plugin, the user must navigate to TrackAnalyzer_Additional_Files, download from `plugins` folder the `JWizardComponent_.jar` and located it into the ImageJ/Fiji `plugins` subfolder. Moreover, to avoid any bugs while running the TraJClassifier motion classification routine, the user must download the .jar files from `jars` folder the .jar files and move them into the ImageJ/Fiji `jars` subfolder. For those users using Fiji, all steps described above can be skipped, the TrackAnalyzer update site can be followed according to the instructions at https://imagej.net/Following_an_update_site.

5.5. Supported Image File Formats

Our plugin deals with a wide range of file formats using Bio-Formats(28), an open-source library from life sciences supporting or reading almost any image format or multidimensional data as z-stacks, time series, or multiplexed images keeping metadata easily accessible. On top of that, the user can access a list of time-lapse images available during the whole procedure to update the analysis as often as needed.

Acknowledgements. We are very grateful to members of the department of Immunology and Oncology at the National Center for Biotechnology for providing us with the time-lapse datasets and their collaborative help in this research.

Funding Statement. This research was supported by the Spanish MICINN (PRE2018-086112) by the FPI fellowship from the Spanish Ministry of Science and Innovation through the Severo Ochoa excellence accreditation SEV-2017-0712-18-1. Also, we would like to acknowledge economic support from Grant PID2019-104757RB-I00 funded by MCIN/AEI/ 10.13039/501100011033/ and "ERDF A way of making Europe", by the "European Union", SEV-2017-0712 funded by MCIN/AEI/10.13039/501100011033, European Union (EU) and Horizon 2020 through grant HighResCells (ERC - 2018 - SyG, Proposal: 810057).

Competing Interests. The authors declare no competing interests.

Data Availability Statement. Source code and documentation for the plugin are available at https://github.com/QuantitativeImageAnalysisUnitCNB/TrackAnalyzer_.

Ethical Standards. The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

Author Contributions. AC, JAGP, and COSS conceived the project and designed the algorithms. AC wrote the software code and performed all experiments. EMGC prepared the samples and acquired the images at the microscope. All authors wrote and revised the manuscript.

References

1. M. Vera, J. Biswas, A. Senecal, R.H. Singer, and H.Y. Park. Single-cell and single-molecule analysis of gene expression regulation. *Annual review of genetics*, 23(50):267–291, 2016.
2. Y. Cui, M. Yu, X. Yao, J. Xing, J. Lin, and X. Li. Single-particle tracking for the quantification of membrane protein dynamics in living plant cells. *Molecular Plant*, 11(11):1315–1327, 2018.
3. R.N.U. Kok, L. Hebert, G. Huelsz-Prince, Y.J. Goos, X. Zheng, K. Bozek, G.J. Stephens, S.J. Tans, and J.S. van Zon. Organoidtracker: Efficient cell tracking using machine learning and manual error correction. *PLoS ONE*, 15(10), 2020.
4. O. Hilsenbeck, M. Schwarzfischer, S. Skylaki, B. Schauberger, P.S. Hoppe, D. Loeffler, K.D. Kokkaliaris, S. Hastreiter, E. Skylaki, A. Filipczyk, M. Strasser, F. Buggenthin, J.S. Feigelman, J. Krumsiek, A.J.J. van den Berg, M. Endele, M. Etzrodt, C. Marr, F.J. Theis, and T. Schroeder. Software tools for single-cell tracking and quantification of cellular and molecular properties. *Nature Biotechnology*, 34:703–706, 2016.
5. C. Manzo and M.F. Garcia-Parajo. A review of progress in single particle tracking: from methods to biophysical insights. *Reports on Progress in Physics*, 78(12):7, 2015.
6. T. Kuhn, J. Hettich, R. Davtyan, and J.C.M Gebhardt. Single molecule tracking and analysis framework including theory-predicted parameter settings. *Scientific Reports*, 11(9465), 2021.
7. K. Jaqaman, D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S.L. Schmid, and G. Danuser. Robust single-particle tracking in live-cell time-lapse sequences. *Nature Methods*, 5(8):695–702, 2008.
8. J.V. Rahm1, S. Malkusch, U. Endesfelder, M.S. Dietz, and M. Heilemann. Diffusion state transitions in single-particle trajectories of met receptor tyrosine kinase measured in live cells. *Frontiers in Computer Science*, 3, 2021.
9. H. Aragaki, K. Ogoh, Y. Kondo, and K. Aoki. Lim tracker: a software package for cell tracking and analysis with advanced interactivity. *Scientific Reports*, 12(2702), 2022.
10. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Tinevez, White D.J., V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona. Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7):676–682, 2012.
11. J. JSchindelin, C.T Rueden, M. C. Hiner, and K. W. Eliceiri. The imagej ecosystem: An open platform for biomedical image analysis. *T. Molecular Reproduction and Development*, 82(7-8):518–529, 2015.
12. J.-Y. Tinevez, N. Perry, J. Schindelin, G. M. Hoopes, G. D. Reynolds, E. Laplantine, and K. W. Eliceiri. Trackmate: An open and extensible platform for single-particle tracking. *Methods*, 115:80–90, 2017.
13. T. Wagner, A. Kroll, CR. Haramagatti, H-G. Lipinski, and M. Wiemann. Classification and segmentation of nanoparticle diffusion trajectories in cellular micro environments. *PLoS ONE*, 12(1), 2017.
14. H. Ewers, A.E. Smith, I.F. Sbalzarini, H. Lilie, P. Koumoutsakos, and A. Helenius. Single-particle tracking of murine polyoma virus-like particles on live cells and artificial membranes. *Proceedings of the National Academy of Sciences*, 102(42), 2005.
15. R. Ferrari, A.J. Manfroi, and W.R. Young. Strongly and weakly self-similar diffusion. *Physica D: Nonlinear Phenomena*, 154(1-2):111–137, 2001.
16. R. Zantl and E. Horn. *Chemotaxis of Slow Migrating Mammalian Cells Analysed by Video Microscopy*. Methods in Molecular Biology (Methods and Protocols). Humana Press, 2011.
17. P. Zengel, A. Nguyen-Hoang, C. Schildhammer, R. Zantl, and V. Kahl. μ-slide chemotaxis: a new chamber for long-term chemotaxis studies. *BMC Molecular and Cell Biology*, 12(21), 2011.
18. EM García-Cuesta, JM. Rodríguez-Frade, SR Gardeta, G D'Agostino, P. Martínez, B. Soler Palacios, G. Cascio, T. Wolf, N. Mateos, R. Ayala-Bueno, CA. Santiago, P. Lucas, L. Llorente, LM. Allende, LI. González-Granado, N. Martín-Cófreces, P. Roda-Navarro, F. Sallusto, F. Sánchez-Madrid, MF. García-Parajo, L. Martínez-Muñoz, and M. Mellado. Altered cxcr4 dynamics at the cell membrane impairs directed cell migration in whim syndrome patients. *Proceedings of the National Academy of Sciences of the United States of America*, 119(21), 2022.
19. L. Martínez-Muñoz, JM. Rodríguez-Frade, R. Barroso, CÓS. Sorzano, JA. Torreño-Pina, CA. Santiago, C. Manzo, P. Lucas, EM. García-Cuesta, E. Gutierrez, L. Barrio, J. Vargas, G. Cascio, YR. Carrasco, F. Sánchez-Madrid, MF. García-Parajo, and M. Mellado. Separating actin-dependent chemokine receptor nanoclustering from dimerization indicates a role for clustering in cxcr4 signaling and function. *Molecular Cell*, 70(1):106–119, 2018.
20. C.O.S Sorzano, L. Martínez-Muñoz, G. Cascio, E.M. García-Cuesta, J. Vargas, M. Mellado, and J.M. Rodriguez Frade. Image processing protocol for the analysis of the diffusion and cluster size of membrane receptors by fluorescence microscopy. *Journal of visualized Experiments: JoVE*, 146, 2019.

21. F. de Chaumont, S. Dallongeville, N. Chenouard, N. Hervé, S. Pop, T. Provoost, V Meas-Yedid, P. Pankajakshan, T. Lecomte, Y. Le Montagner, T. Lagache, A. Dufour, and J.C. Olivo-Marin. Icy: an open bioimage informatics platform for extended reproducible research. *Nature Methods*, 9:690–696, 2012.
22. N. Chenouard, I. Bloch, and J.C. Olivo Marin. Multiple hypothesis tracking for cluttered biological image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(11):2736–3750, 2013.
23. Marloes Arts, Ihor Smal, Maarten W. Paul, Claire Wyman, and Erik Meijering. Particle mobility analysis using deep learning and the moment scaling spectrum. *Scientific Reports*, 9(1), November 2019.
24. Anthony R. Vega, Spencer A. Freeman, Sergio Grinstein, and Khuloud Jaqaman. Multistep track segmentation and motion classification for transient mobility analysis. *Biophysical Journal*, 114(5):1018–1025, March 2018.
25. J.P. Siebrasse, I. Djuric, U. Schulze, M.A. Schlüter, H. Pavenstädt, T. Weide, and U. Kubitscheck. Trajectories and single-particle tracking data of intracellular vesicles loaded with either snap-crb3a or snap-crb3b. *Data in Brief*, 7(9465):1665–1669, 2016.
26. S. Sternberg. Biomedical image processing. *Computer*, 16(1):22–34, 1983.
27. J. desRivieres and J. Wiegand. Eclipse: A platform for integrating development tools. *IBM Systems Journal*, 43(2):371–383, 2004.
28. M. Linkert, C.T. Rueden, C. Allan, J.M. Burel, W. Moore, A. Patterson, B. Loranger, J. Moore, C. Neves, D. Macdonald, A. Tarkowska, C. Sticco, E. Hill, M. Rossner, K.W. Eliceiri, and J.R. Swedlow. Metadata matters: access to image data in the real world. *The Journal of Cell Biology*, 189(5):777–82, 2010.

Appendix C

Real-Time Correction of Chromatic Aberration in Optical Fluorescence Microscopy

López AC, Conesa P, Oña Blanco AM, Gómez-Pedrero JA, Sorzano COS. Real-time correction of chromatic aberration in optical fluorescence microscopy. *Methods Appl Fluoresc*. 2023 Jul 3;11(4). doi: 10.1088/2050-6120/ACE153. PMID: 37352866.

IOP Publishing Methods Appl. Fluoresc. 11 (2023) 045001 https://doi.org/10.1088/2050-6120/ace153

Methods and Applications in Fluorescence

RECEIVED
20 April 2023

REVISED
15 June 2023

ACCEPTED FOR PUBLICATION
23 June 2023

PUBLISHED
3 July 2023

PAPER

Real-time correction of chromatic aberration in optical fluorescence microscopy

Ana Cayuela López[1,4], **Pablo Conesa**[1,4], **Ana María Oña Blanco**[2], **José Antonio Gómez-Pedrero**[3] and **Carlos Oscar S Sorzano**[1]

[1] Biocomputing Unit, National Centre for Biotechnology, 28049 Cantoblanco, Madrid, Spain
[2] Advanced Light Microscopy Unit, National Centre for Biotechnology, 28049 Cantoblanco, Madrid, Spain
[3] Applied Optics Complutense Group, Faculty of Optics and Optometry, University Complutense of Madrid, 28037 Madrid, Spain
[4] These authors contributed equally to this work.

E-mail: acayuela@cnb.csic.es, pconesa@cnb.csic.es, aoblanco@cnb.csic.es, jagomezp@fis.ucm.es and coss@cnb.csic.es

Keywords: elastic registration, chromatic aberration, B-splines, Scipion, shift

Abstract

Multi-color fluorescence imaging is a powerful tool for studying the spatial relationships and interactions among sub-cellular structures in biological specimens. However, if improperly corrected, geometrical distortions caused by mechanical drift, refractive index mismatch, or chromatic aberration can lead to lower image resolution. In this paper, we present an extension of the image processing framework of Scipion by integrating a protocol called OFM Corrector, which corrects geometrical distortions in real-time using a B-spline-based elastic continuous registration technique. Our proposal provides a simple strategy to overcome chromatic aberration by digitally re-aligning color channels in multi-color fluorescence microscopy images, even in 3D or time. Our method relies on a geometrical calibration, which we do with fluorescent beads excited by different wavelengths of light and subsequently registered to get the elastic warp as a reference to correct chromatic shift. Our software is freely available with a user-friendly GUI and can be broadly used for various biological imaging problems. The paper presents a valuable tool for researchers working in light microscopy facilities.

Abbreviations

The following abbreviations are used in this manuscript:

CA	Chromatic Aberration
ACA	Axial Chromatic Aberration
LCA	Lateral Chromatic Aberration
TIRF	Total Internal Reflection Microscopy
GUI	Graphical User Interface
IQR	Interquartile Range
FOV	Field of View
RBF	Radial Basis Functions

1. Introduction

Over the past few years, many technological advancements have been made in single molecule-based super-resolution microscopy techniques [1]. One of the imaging modalities in fluorescence microscopy is multi-color fluorescence imaging, which enables the differentiation of proteins and structures of interest in both living and fixed cells [2]. This technique also helps to determine the intracellular relationships or interactions between sub-cellular structures [1]. However, mechanical drift, chromatic aberrations caused by optical elements, refractive index mismatching between the objective and immersion medium, and dispersion in biological samples can lead to decreased image resolution [3]. In all optical systems, chromatic aberration (CA) occurs due to differences in the refractive index among optical components, causing the light wavelengths to focus at slightly different angles. The phenomenon is noticeable in the acquired images because the color channels are misaligned, causing colored fringes at the edges and high-contrast regions [4]. This can significantly decrease the image quality [5, 6]. In the case of biological applications, CA may negatively affect multi-channel studies of dynamic processes in cells, such as colocalization research and object-based analysis. Therefore, it is essential to quantify and correct these aberrations properly. In this sense, we can find two types of CA: lateral, which occurs along the x-y axis, and axial, along the z-axis. Axial chromatic aberration (ACA) comes when lenses through the optics have a refractive index that varies with the wavelengths, focal distance, and image magnification. This directly affects the image focus and resolution, giving rise to image blur [7, 8]. By contrast, lateral chromatic aberration (LCA) occurs when there is variation in the magnification of different light colors. This prompts protruding image edges and deviations when two color images are superimposed.

Total Internal Reflection Fluorescence (TIRF) microscopy is a powerful optical technique to selectively acquire images of molecules in an aqueous environment with a high refractive index [9]. This approach provides extremely thin axial optical sectioning with a high signal-to-noise ratio allowing microscopists to image fluorescent membrane-associated events in living cells (cell adhesion, hormone binding, molecule transport, exocytotic and endocytotic processes, ...) as well as molecules located at the medium interface with a higher refractive index and a lower at an incidence angle bigger than critical angle [10]. The optical system may be either prism-based or objective-based to reach total internal reflection to optimize each color uniquely and independently, enabling the imaging of multiple colors simultaneously. In the first approach, a prism is attached to the coverslip's surface, which directs a focused light beam or laser toward the medium interface at the critical angle. The objective-based approach, instead, is the system mainly used, and the light is directed to the specimen through the objective, which simultaneously collects the emitted fluorescence light. In this context, dealing with multi-color and multi-angle TIRF may result challenging, and unfortunately, LCA is essentially inherent. This LCA induces shifts, rotations, and scaling differences among channels.

The Advanced Microscopy Facility of our institute has a TIRF microscope with a W-VIEW Gemini system from Hamamatsu [11], an image-splitting optics device. It was adjusted to split the signal on the camera chip by wavelength in two channels (two pairs of images) with a dichroic mirror. This optical component allows high-speed acquisition with a vast variety of fluorescence applications and permits simultaneous two-wavelength (dual) imaging by one camera due to its optical design. Additionally, this system integrates a mechanism to compensate ACA and LCA. This mechanism is based on a correction lens unit in the long wavelength path, and it can improve the magnification difference of two wavelength images caused by LCA. Furthermore, this system was designed to be easily adjusted with a camera due to integrating a fast and straightforward alignment mechanism to realign the optics. Besides that, the Gemini system has a feature to control temperature and time stability, hence ensuring the alignment consistency of two channels over time for dual-wavelength imaging. Despite these ideal specifications, this optical component could not properly overcome the effect of LCA in our TIRF infrastructure. Interestingly, this instability is explicitly mentioned in the microscope documentation [11].

The commercial company proposed several commercial solutions (hardware- and software-based) to solve this undesirable misalignment effect. Still, most of these approaches were expensive, not intuitive at the user level, non-effective in covering the full FOV, and time-consuming. CA correction has also been addressed in the scientific literature [12]. The correction methods can be grouped into hardware- and software-based [7]. Within hardware-based methods, apochromatic lenses are developed to set into focus in the same plane, red, green, and blue wavelengths. Also, active lens control systems are designed [13] to correct CA by adjusting the distance between the image plane and lens. Nevertheless, apochromatic lenses are affected by residual errors too big to be ignored, and the lens control system requires prior knowledge of the magnification and image shift degree. On the other hand, most software-based methods to compensate CA are based on image registration [14], false color techniques [15], and post-demosaicking correction based on pixel re-sampling and high-pass replication [16]. However, none of these methods are readily available and easily integrated within the standard procedures of a microscopy facility.

In this paper, we use image registration to compensate the geometrical distortions induced by LCA. This technique works by spatially registering images such that corresponding features are consistent in geometry. It involves identifying corresponding features or pixels in two or more images and then applying a geometric transformation to align them. The transformation can be rigid, affine, or non-rigid, depending on the type and degree of misalignment. This paper uses B-spline-based elastic image registration [17] for modeling deformations in biological imaging problems [18]. This technique has several advantages, such as coping with a wide range of deformations, including non-linear [17]. The registration process is based on image similarity, deformation consistency, and cubic B-spline regularization [19]. This technique ensures high-quality interpolation of the images and allows an arbitrarily fine representation of the deformation field by reducing the spacing among splines. B-spline-based elastic image registration is advantageous in many biological imaging problems, such as tracking the movement of cells or analyzing the shape of tissues. Accordingly, B-spline based methods have gained popularity in image registration due to their flexibility and ability to accurately capture complex deformations. Among its advantages, this approach allows for localized control over the deformation field by dividing the image into smaller regions (control points) thus, representing localized deformations more effectively. This can be advantageous when dealing with complex and non-uniform deformations, which can vary spatially across the image. Moreover, B-spline method provides smooth and continuous deformations, something particularly relevant in microscopy, where smooth deformations are desirable to preserve anatomical structures or avoid introducing artifacts. B-spline methods are computationally attractive, which is especially worthy when dealing with large image datasets or real-time approaches. On the other hand, alternative methods such as radial basis functions (RBF) with compact support, also provide accurate deformations, particularly for small and localized deformations with high precision. While RBF methods are widely used for their interpolation properties which can capture fine details exactly, they are still challenging handling large deformations or global registration tasks due to the compact support limitation. Therefore, the choice of the most suitable registration method over alternative approaches depends on the specific requirements of the application, the study context, the nature of the deformations and the available computational resources. For such, the B-spline-based method was selected for this study based on its ability to handle complex deformations, provide smooth results as well as efficient computation.

Our method utilizes multispectral fluorescent beads as a reference for image registration and drift correction [20]. These fluorescent beads are excited by different wavelengths of light and emit differently in the same wavelength range as the applied dyes. The shift between image channels is recorded and registered for the warp transformation to correct further chromatic shifts in images acquired under the same imaging setup. This elastic registration process involves finding the image transformation that can best map one image into the other. The integrated algorithm extends the elastic (non-linear) registration approach [17] by providing an almost invertible deformation field, allowing bidirectional registration. This ensures that source image A can be mapped onto target image B and vice versa in a single computation, thereby reducing the optimizer likelihood of being trapped in a local minimum and enabling simultaneous registration of any number of images.

A requirement for facilities is that the solution must work in real-time while the TIRF videos are acquired. This way, the user can bring home the CA-corrected data after finishing the microscopy session. To this end, we have developed a protocol called *ofm-correction - OFM Corrector* based on the bUnwarpJ [19, 21] plugin (available under ImageJ [22] or Fiji [23] distribution) and integrated into the Scipion's image processing framework [24]. Before acquiring the TIRF videos, the microscope operator must calibrate the deformation field for that particular acquisition (because the deformation field depends on the ambient temperature, the specific magnification setup, and the Field of View (FOV) region being imaged). Once calibrated, the deformation field is used to correct all the videos acquired with the same conditions. Our protocol offers a unified graphical user interface (GUI), package interoperability, a simple and cost-effective strategy to overcome geometrical distortions, and workflow monitoring for the streaming registration process (see figure 1). Our software is freely available within Scipion framework and can be used in any microscopy setup with geometric distortions affecting the acquired videos.

2. Real-time correction of geometric distortions

This section describes the algorithmic approach to solving geometric distortions. We first introduce the procedure to measure the geometric distortions at the microscope experimentally. Then, we describe our algorithm to construct a mathematical description of the deformation field and correct it. Finally, we present the framework that allows real-time correction with images in streaming.

2.1. Experimental measure of the deformation field

A possible way to experimentally measure the deformation field is by recording images of known objects. Multi-spectral fluorescent beads are suitable for this purpose because they fluoresce at various wavelengths, and any image misalignment can be easily detected [25]. Figure 2 shows the conceptual setup from one of our experiments. The various wavelengths are projected differently by the dichroic mirror, so the image of the same bead is projected at two different locations. From a pair of these images, we can estimate the relative deformation field ($\mathbf{g}^{12}(\mathbf{s})$ in the equation (1) below).

As shown in figure 2(D), we can see that the LCA shift depends on the region of the FOV being imaged and the lateral distance to the center of the image.

2.2. Elastic image registration

Let us consider a pair of images acquired in Channels 1 and 2, $I^1(\mathbf{s})$ and $I^2(\mathbf{s})$, where $\mathbf{s} = (x, y)$ is a 2D vector with the pixel coordinate. Elastic image registration assumes that there is a deformation field, $\mathbf{g}^{12}$, that transforms coordinates from one channel onto the coordinates of the other:

$$I^1(\mathbf{s}) = I^2(\mathbf{s} + \mathbf{g}^{12}(\mathbf{s})) \quad (1)$$

In case there is no geometrical distortion, then $\mathbf{g}^{12}(\mathbf{s}) = \mathbf{0}$ for all $\mathbf{s}$, and the two channels should superpose exactly. However, if they do not, we look for the deformation field that minimizes the error between these two images. To estimate the deformation field, it is important to use objects whose emission in both channels is the same (see the previous section).

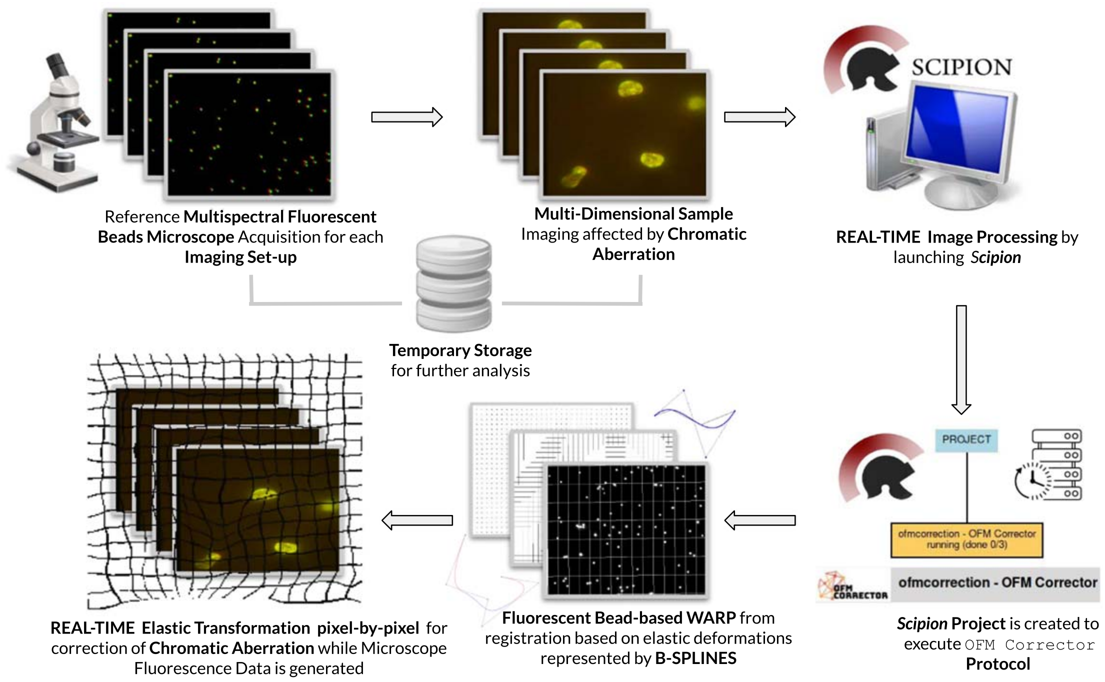

Figure 1. Graphical representation of the workflow to reach real-time correction of geometric misalignment among channels in multi-dimensional images acquired with fluorescence microscopy using multi-spectral fluorescent beads through Scipion software.

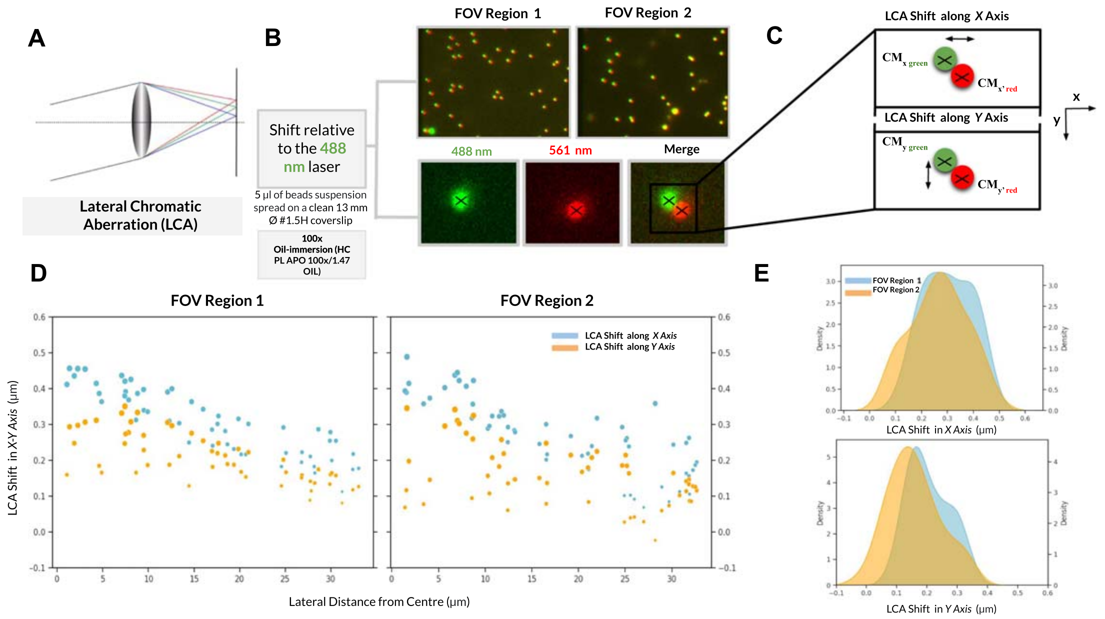

Figure 2. Measuring LCA shifts among channels with fluorescence beads images by identification of the x-y center of mass. (A) Schematic illustration of lateral chromatic aberration (LCA). (B) Illustration of the different geometric locations of the same bead imaged at two different wavelengths. (C) Measurement of LCA. Centroid positions for each channel are shown by black crosses. (D) LCA (at 561 nm) shifts relative to 488 nm for a hundred beads at two different regions of the FOV. For all cases, the shifts along the x-y axes for all beads are small (less than 1 μm). Dot size increases according to LCA shift. (E) Probability density function (via Kernel density estimation) of LCA Shifts along X-Y axes (μm) at two different regions of the FOV.

We use a B-splines parametrization of the deformation field as explained in [17]. This way, the deformation field is calculated as

$$\mathbf{g}^{12}(\mathbf{s}) = \sum_{ij} \mathbf{c}_{ij}^{12} B\left(\frac{x - ih}{h}\right) B\left(\frac{y - jh}{h}\right) \quad (2)$$

where i and j are indexes over a regular lattice of B-spline functions, B, whose separation between grid points is h is both directions. The coefficients $\mathbf{c}_{ij}^{12} \in \mathbb{R}^2$ are the ones that control the amount of deformation in x and y.

The transformation $\mathbf{g}^{12}(\mathbf{s})$ may not be invertible. It has been observed that this deformation field is better

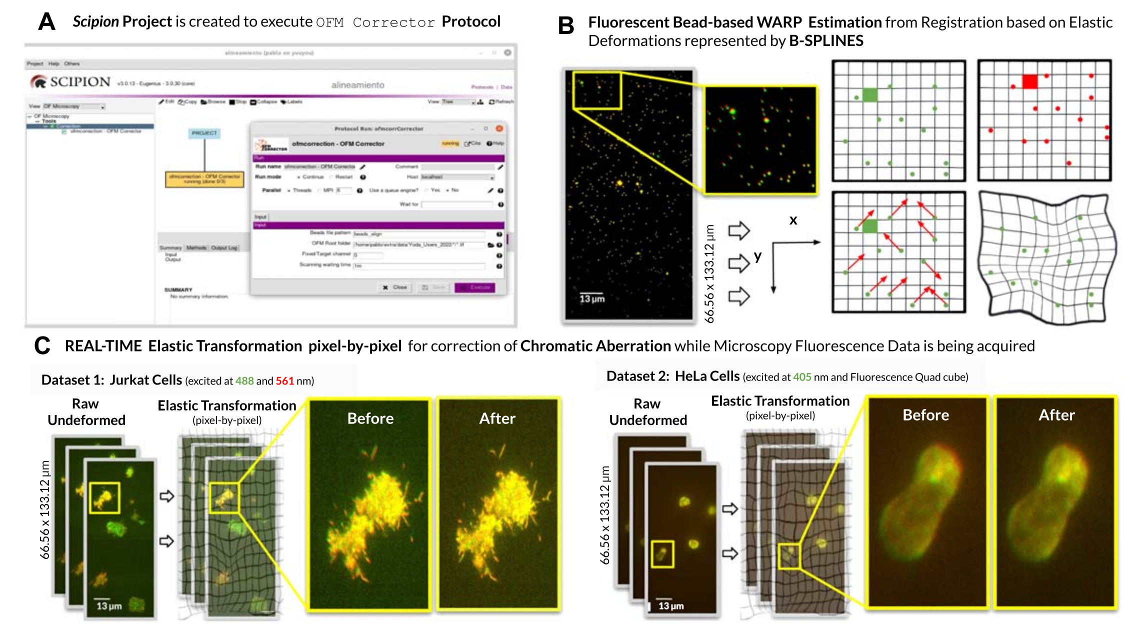

Figure 3. Schematic overview of using Scipion to compensate LCA shifts in volumetric biological images. (A) Scipion Graphical User Interface. (B) Once the protocol is launched, the deformation field is estimated as described in section 2.2. (C) The deformation field is corrected for all the input videos corresponding to biological samples.

estimated if the deformation is computed bidirectionally:

$$I^1(\mathbf{s}) = I^2(\mathbf{s} + \mathbf{g}^{12}(\mathbf{s}))$$
$$I^2(\mathbf{s}) = I^1(\mathbf{s} + \mathbf{g}^{21}(\mathbf{s})) \quad (3)$$

and $\mathbf{g}^{12}$ and $\mathbf{g}^{21}$ are supposed to be approximate inverses of each other:

$$\mathbf{g}^{21}(\mathbf{g}^{12}(\mathbf{s})) \approx \mathbf{s} \quad (4)$$

This approximate inverse condition is called a consistency constraint.

The $\mathbf{c}_{ij}^{12}$ and $\mathbf{c}_{ij}^{21}$ coefficients are determined by minimizing the following error function

$$\begin{aligned} E = {} & \sum_{\mathbf{s}} \|I^1(\mathbf{s}) - I^2(\mathbf{s} + \mathbf{g}^{12}(\mathbf{s}))\|^2 + \sum_{\mathbf{s}} \|I^2(\mathbf{s}) \\ & -I^1(\mathbf{s} + \mathbf{g}^{21}(\mathbf{s}))\|^2 \qquad \text{image dissimilarity} \\ & +w_c \sum_{\mathbf{s}} \|\mathbf{g}^{21}(\mathbf{g}^{12}(\mathbf{s})) - \mathbf{s}\|^2 \\ & \text{consistency error} \\ & +w_d \left(\sum_{\mathbf{s}} (\|\nabla \text{div} \mathbf{g}^{12}(\mathbf{s})\|^2 + \|\nabla \text{div} \mathbf{g}^{21}(\mathbf{s})\|^2) \right) \\ & \text{regularization divergence} \\ & +w_r \left(\sum_{\mathbf{s}} (\|\nabla \text{rot} \mathbf{g}^{12}(\mathbf{s})\|^2 + \|\nabla \text{rot} \mathbf{g}^{21}(\mathbf{s})\|^2) \right) \\ & \text{regularization rotational} \end{aligned} \quad (5)$$

where w_c, w_d, and w_r are weights that control the relative weight of the different terms The minimization of the error term with respect to the B-spline coefficients was explained in [19, 21], and it is publicly available through the BUnwarpJ plugin of ImageJ. We have not reimplemented this algorithm, but we call it through Fiji.

We estimate the deformation field, which is encoded through the $\mathbf{c}_{ij}^{12}$ and $\mathbf{c}_{ij}^{21}$ coefficients, during the calibration step. These coefficients are saved after calibration and reused to produce aligned images. For instance, to correct the image from Channel 2 so that it is registered with Channel 1, we construct the image $(I^2)'$ as

$$(I_t^2)'(\mathbf{s}) = I_t^2(\mathbf{s} - \mathbf{g}^{12}(\mathbf{s})) \quad (6)$$

The subindex t has been introduced to represent the different time frames within a video. The distortion correction above is applied to all the video frames acquired by the TIRF microscope.

2.3. Geometric corrections in real-time

Microscopy facilities continuously receive users acquiring their images on the samples of their interest. In this scenario, it is essential for facilities to keep up to the highest quality standard. Having a microscope with severe geometrical distortions, such as the one presented in this paper, is a drawback for the facility. Therefore, we have integrated the elastic registration algorithm described above into an image-processing workflow engine called Scipion [24]. This workflow engine is also developed by our laboratory. This engine allows image processing in streaming [26]: the newly acquired images are geometrically corrected as soon as they are written in their folder (see figure 3). In this way, the user can bring home the already corrected images. The plugin is called *ofmcorrection* and the protocol *OFM Corrector*. The beads images are one of the inputs of the protocol. The protocol first estimates the deformation field to correct. Then it applies it to all videos in the input folder (it must be noted that this correction is applied on-the-fly, so new videos can arrive once the protocol has started). The geometrical distortion corrected is the same for all videos and is measured at the beginning of the acquisition as described in section 2.1. This calibration image is only valid for experimental images acquired under the same exposure time, camera gain, region of the field of view, brightness, photo-stability, temperature, and ambient light in the facility.

3. Materials and methods

3.1. Code availability

OFM Corrector protocol was written in Groovy and integrated into Scipion framework by using Python. The complete source code of the algorithm integrated in Scipion software is available at https://github.com/acayuelalopez/bUnwarpJ_code. Our protocol can be used for real-time processing within the Scipion framework. You can install it by using the Scipion software following the Scipion's installation guide (https://scipion-em.github.io/docs/release-3.0.0/docs/scipion-modes/how-to-install.html).

3.2. Experimental methods

3.2.1. Calibration sample preparation

Multispectral fluorescent beads suspension (TetraSpeck^{TM} Microspheres, 0.2 μm, fluorescent blue/green/orange/dark red, TermoFisher Scientific) were used as a reference for calibration of the image alignment. Five microlitres of the beads suspension were pipetted and spread on a clean 13 mm Ø #1.5H coverslip (Menzel Gläser) for 1 hour for adhesion to the glass. The coverslip was placed on a slide and sealed with enamel. Once the sample beads were prepared, they were placed on the immersion oil objective (Leica^{TM} Immersion Oil) of the TIRF microscope to acquire several images. The fluorescent beads were simultaneously excited by two different wavelengths of light: 488, and 561 nm lasers. In our acquisition, we used a 100x oil-immersion objective (HC PL APO 100x/1.47 OIL) with a Leica DMi8 S with TIRF module microscope equipped with Hamamatsu Flash 4 digital sCMOS camera.

3.2.2. Biological sample preparation

Jurkat cells (American Type Culture Collection, ATCC TIB-152) or HeLa cells (ATCC CCL-2) were maintained in culture using a complete growth medium (RP 1640 or DMEM, Gibco, plus 10% fetal calf serum) at 37 °C and 5% CO_2.

Jurkat cells were transiently transfected with plasmids to express cell membrane receptors fused to EGFP or mCherry reporters using a BioRad electroporator (2 × 104 cells in RP 1640 with 10% fetal calf serum. 280V, 975 mF) and imaging 24 hours later.

DNA Hela cells were stained 10 minutes with Hoechst 33342 (ThermoFisher Scientific) at 0.5 mg/mL prepared in cell culture medium. The staining solution was gently removed and the cells were washed with culture medium to remove unbound dye. To proceed with image acquisition, the cells were left in DMEM medium plus 10% fetal calf serum. For the imaging experiments, the cells were seeded in a μ-Dish 35 mm, high glass bottom dish (Ibidi) at a density of 20,000 to 50,000 cells per well.

3.2.3. Dual TIRF imaging

Dual TIRF experiments were performed using a Leica DMi8 S with a TIRF module microscope equipped with a Hamamatsu Flash 4 digital sCMOS camera (Hamamatsu), a 100x oil-immersion objective (HC PL APO 100x/1.47 OIL), and the 405, 488 and 561 nm laser lines for the illumination of the samples. Two types of W-View Gemini splitter (Hamamatsu) were used for simultaneous image acquisition. This image-splitting optic provides a pair of dual-wavelength or polarization images separated by a dichroic mirror in a single camera. The beam divider allowed us to obtain two separate sample images of the same field of view (FOV) on the same camera chip. When the W-View optics were in place, the Hamamatsu W-View Gemini option was activated in the software interface, and the images of the simultaneously imaged beads were manually aligned to correct the focus, zoom, and x and y shift between the two channels using EPI laser position (without laser penetration depth). Once manually aligned, the reference images were acquired with an EPI laser position. Then, the images were acquired with the biological sample under the acquisition conditions required by the experimental design. The microscope was equipped with an incubator and temperature control units; experiments were performed at 37 °C with 5% CO2. Z-stabilization was ensured by the adaptive focus control (AFC) on the microscope.

Image sequences (500 frames for Jurkat cells and 11 frames for HeLa cells) were acquired with a 90 ms/frame rate for Jurkat cells and 60 ms/frame for HeLa cells. The penetration depth of the evanescent wave was 90 nm for Jurkat cells and EPI laser position for HeLa cells. The images have 512 × 1024 (0.13 × 0.13 μm pixel size) for Jurkat cells or 1024 × 2048 (0.065 × 0.065 μm pixel size) pixels for HeLa cells and were acquired at 16-bit.

For EGFP/mCherry imaging, the 488 and 561 nm excitation lasers lines were used simultaneously, and the fluorescence Dual cube (GFP/Ch-T) employed has excitation filters between 483-493 and 550-568 nm, an emission filters between 507-553 and LP575 nm and dichromatic beamsplitters at 500 and 575 nm. The bandpass filters used were for GFP/DsRED dual-band imaging set (FF01-512/25-25, FF01-630/92-25, and dichroic mirror FF560-FDi01-25 × 36).

For Hoechst 33342 imaging, the 405 excitation laser was used with the fluorescence Qua-T cube. The

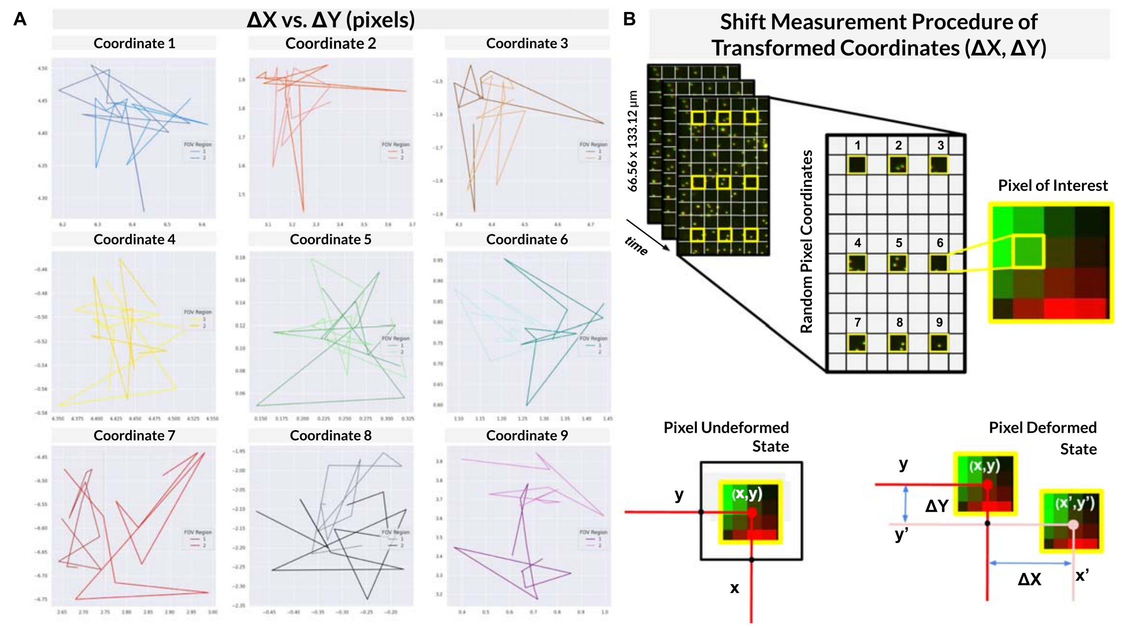

Figure 4. Stability of the deformation field over time. (A) For some representative coordinates, we show the trajectory over time of the deformation field. Displacements are expressed in pixels (the pixel size is 0.065 μm. (B) Schematic visualization of the shift measurement procedure. x, y are the pixel coordinates (undeformed state) and x', y' (deformed state) are the pixel coordinates after elastic transformation, and ΔX, ΔX are the displacement among them.

Qua-T cube has an excitation filter between 397–413 nm, an emission filter between 420–480 and 500–550 nm, and dichromatic beamsplitters at 415 and 495 nm, respectively, among others.

4. Results

We started by estimating and characterizing the deformation field and its stability over time. Then, we applied our method to two biological data sets consisting of dual-wavelength images acquired with two different z-axis depths (described in detail in section 3.2.2).

4.1. Stability of the deformation field over time

We acquired images of the beads exited with two laser channels, 488 and 561 nm., every hour during the day (12 measures in total). We focused the microscope on two different regions of the FOV. We estimated the deformation field and evaluated its stability in every location. Figure 4 shows the trajectories over time of the x and y components of $\mathbf{g}^{12}(\mathbf{s})$ (figure 4(A)) for 9 points uniformly distributed over the region being imaged (figure 4(B)).

Depending on the FOV region and the location within this region, the deformation field can be as high as 6.6 pixels (about 0.4 μm., figure 4(A)), or even higher for pixels closer to the image border. In the center of the region, figure 4(E), the deformation is relatively small. For this reason, that was the only area that the microscope users used to analyze before the facility incorporated our real-time correction. To facilitate the visualization over the whole region, in figure 5, we show the mean and standard deviation of the deformation field over time at every location for two different regions. Interestingly, the vertical and horizontal distortions behave differently. This is due to the polarization of the light being used and the possible anisotropic nature of the crystals of which the different optical devices along the path are made.

4.1.1. Dataset 1: jurkat cells

This dataset consists of four series containing 1000 planes with a voxel size of 0.13 $\times$ 0.13 $\times$ 1 μm^3 saved as Leica File Format (figure 3(C), top). Jurkat cells were transiently transfected, expressing cell membrane receptors fused to EGFP (green, 488 nm.) or mCherry (red, 561 nm.) reporters. Two simultaneously acquired image sequences (500 frames) GFP/DsRED dual-band imaging filters were used. The sample was illuminated with 488 and 561 nm. laser lines and fluorescence dual cube was used. The images were acquired with a frame rate of 90 ms/frame, and the penetration depth of the evanescent wave was 90 nm. The images were acquired with a pixel map of 512 $\times$ 1024 pixels (66.56 $\times$ 133.12 μm) and a bit-depth of 16 bits. As shown in figure 3(C), the raw images are heavily affected by LCA. After applying the bead-based warping transformation, LCA misalignment was always fully corrected (500 frames).

4.1.2. Dataset 2: HeLa cells

This dataset consists of 6 series containing 22 planes with a voxel size of 0.13 $\times$ 0.13 $\times$ 1 μm^3 saved as Leica File Format (figure 3(C), bottom). The nuclei of HeLa cells were stained with Hoechst 33342. Two

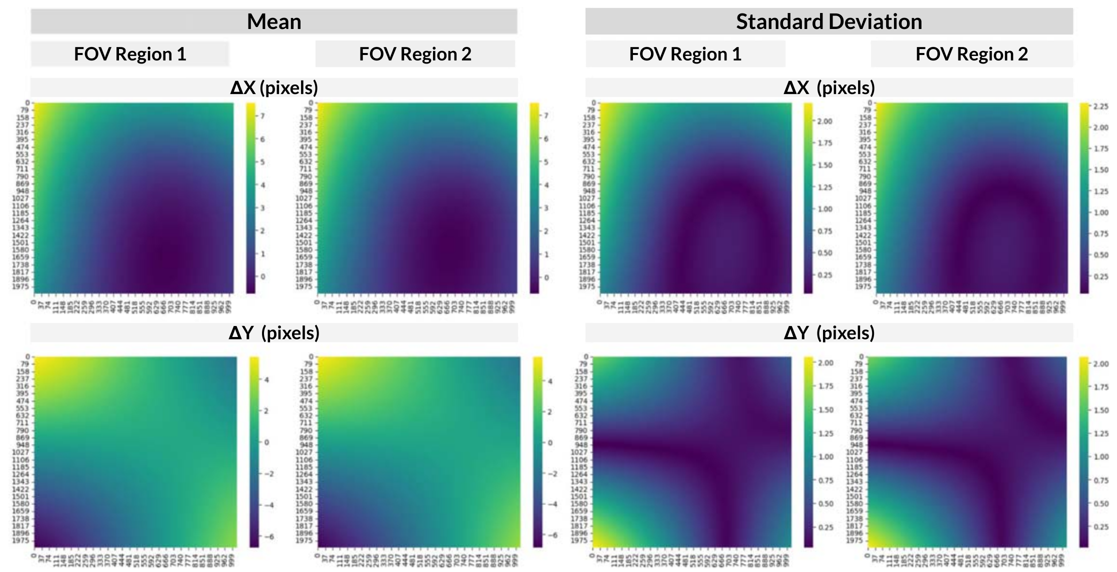

Figure 5. Mean and standard deviation of the deformation field over time at two different regions of the FOV. The size of the acquired region is 1024 × 2050 pixels (66.56 × 133.12 μm.)

simultaneously acquired image sequences (11 frames) light polarization filters set were used. The sample was illuminated with a 405 nm. laser line and a fluorescence Quad cube was used. The images were acquired with a 60 ms/frame rate and EPI laser position. The pixel size and bit-depth of the acquired images were the same as in the previous dataset. As seen in figure 3(C), the misalignment caused by LCA was also corrected following the same procedure as in the previous dataset.

5. Discussion and conclusions

Chromatic aberration is a prevalent issue in multicolor imaging. However, geometric distortions may appear for other experimental reasons, such as imperfections of the optical elements, the mismatch between the refractive index of the objective and immersion medium, or differential dispersion inside the biological samples. We have developed an inexpensive and very efficient solution to a problem that the company commercializing the microscope could not solve with a more accurate physical construction of the dichroic mirror. This problem severely limited the region the microscope users could analyze in their biological experiments. Our software solution is integrated into a protocol called the *OFM Corrector*, freely accessible within the Scipion framework. Scipion offers the possibility of applying the geometrical correction in streaming and real-time, providing almost instant aberration-corrected images. This way, our solution favorably compares to expensive optical solutions. Additionally, it is much more general as it does not only address chromatic aberration but any other source of geometrical distortions. This method can be applied easily to all future acquisitions in the light microscopy facility by capturing a reference calibration image for each condition. This calibration step depends on the specific imaging set-up, including excitation laser lines, objective lens, temperature stability, and exposure time. Also, selecting the most suitable multi-spectral fluorescent beads based on their signal and size is vital to ensure that the bead diameter is reasonably above the microscope's resolution, providing a sufficient signal-to-noise ratio. Our solution is not limited to correcting geometrical distortions between two channels. Any number of channels can be simultaneously corrected. One of the channels must act as the reference channel, while all the others are corrected to match the reference.

Overall, this protocol has the potential to be widely adopted in light microscopy facilities and carries significant implications for the field of biological imaging. Future efforts may concentrate on expanding the protocol's capabilities to address additional optical distortions found in biological imaging.

Acknowledgments

We would like to thank the members from the Advanced Light Microscopy facility at the National Center for Biotechnology, Gianluca DÁgostino and Jaime Fernández de Córdoba for providing us with the dataset of cells acquired by dual TIRF microscopy and also for their collaborative help in this research.

Author contributions

ACL, PCM, JAGP, and COSS conceived the project and designed the algorithms. ACL and PCM wrote the software code and performed all experiments. AOB prepared the samples and acquired the images at the microscope. All authors have read and agreed to the published version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by the Spanish MICINN (PRE2018-086112) by the FPI fellowship from the Spanish Ministry of Science and Innovation through the Severo Ochoa excellence accreditation SEV-2017-0712-18-1. Also, we would like to acknowledge economic support from Grant PID2019-104757RB-I00 funded by MCIN/AEI/ 10.13039/501100011033/ and 'ERDF A way of making Europe', by the 'European Union', SEV-2017-0712 funded by MCIN/AEI/ 10.13039/501100011033, European Union (EU) and Horizon 2020 through grant HighResCells (ERC - 2018 - SyG, Proposal: 810057).

Conflicts of interest

The authors declare no conflict of interest.

ORCID iDs

Ana Cayuela López https://orcid.org/0000-0001-7728-9867
José Antonio Gómez-Pedrero https://orcid.org/0000-0001-7043-3256

References

[1] Tam J, Cordier G A, Borbely J S, Sandoval Álvarez Á and Lakadamyali M 2014 Cross-talk-free multi-color storm imaging using a single fluorophore *PLoS One* **9** 7
[2] Dean K and Palmer A 2014 Advances in fluorescence labeling strategies for dynamic cellular imaging *Nat. Chem. Biol.* **10** 7
[3] Matsuda A, Schermelleh L, Hirano Y, Haraguchi T and Hiraoka Y 2018 Accurate and fiducial-marker-free correction for three-dimensional chromatic shift in biological fluorescence microscopy *Sci. Rep.* **8** 1
[4] Rudakova V and Monasse P 2014 Precise correction of lateral chromatic aberration in images *Image and Video Technology* ed R Klette *et al* (Berlin Heidelberg: Springer Berlin Heidelberg) pp 12–22
[5] Yu H, Zhou Z, Qi Y, Zhang X and Wei Q H 2019 Pancharatnam-berry optical lenses *J. Opt. Soc. Am.* B **36** 107–11
[6] Yousefzadeh C, Jamali A, McGinty C and Bos P J 2018 Achromatic limits of pancharatnam phase lenses *Appl. Opt.* **57** 1151–8
[7] Huang J, Xue Q, Wang Z and Gao J 2016 Analysis and compensation for lateral chromatic aberration in a color coding structured light 3d measurement system *Sensors* **16** 9
[8] Leiwe M N, Fujimoto S and Imai T 2021 Post hoc correction of chromatic aberrations in large-scale volumetric images in confocal microscopy *Frontiers in Neuroanatomy* **15** 760063
[9] Axelrod D 2001 Total internal reflection fluorescence microscopy in cell biology *Traffic* **2** 764–74
[10] Fish K N 2009 Total internal reflection fluorescence (TIRF) microscopy *Current Protocols in Cytometry* **50** 12.18.1–12.18.13
[11] Hamamatsu Photonics K K W-VIEW GEMINI Image Splitting Optics A12801-01 Technical Note, https://hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/sys/SBIS0101E_A1280101.pdf
[12] Roberts D, Liao Z, Hwang J Y, Nersisyan S R, Tabirian N, Steeves D M, Kimball B R and Bunning T J 2018 Chromatic aberration corrected switchable optical systems *Liquid Crystals XXII* ed I C Khoo (International Society for Optics and Photonics, SPIE) vol 10735 107350Q
[13] Willson R G and Shafer S A 1991 Active lens control for high precision computer imaging *Proceedings. 1991 IEEE International Conference on Robotics and Automation, vol 3* pp 2063–70
[14] Lluis-Gomez A and Edirisinghe E A 2012 Chromatic aberration correction in raw domain for image quality enhancement in image sensor processors *2012 IEEE VIII International Conference on Intelligent Computer Communication and Processing* pp 241–4
[15] Chang J, Kang H and Kang M G 2013 Correction of axial and lateral chromatic aberration with false color filtering *IEEE Trans. Image Process.* **22** 1186–98
[16] Korneliussen J T and Hirakawa K 2014 Camera processing with chromatic aberration *IEEE Trans. Image Process.* **23** 4539–52
[17] Sorzano C O S, Thévenaz P and Unser M 2005 Elastic registration of biological images using vector-spline regularization *IEEE Trans. Biomed. Eng.* **52** 652–63
[18] Uji A, Ooto S, Hangai M, Arichika S and Yoshimura N 2013 Image quality improvement in adaptive optics scanning laser ophthalmoscopy assisted capillary visualization using b-spline-based elastic image registration *PLoS One* **8** 11
[19] Arganda-Carreras I, Sorzano C O S, Marabini R, Carazo J M, Ortiz-de C, Solorzano and Kybic J 2006 Consistent and elastic registration of histological sections using vector-spline regularization *Lect. Notes Comput. Sci.* **4241** 85–95
[20] Erdelyi M, Rees E, Metcalf D, Kaminski Schierle G S, Dudas L, Sinko J, Knight A E and Kaminski C F 2013 Correcting chromatic offset in multicolor super-resolution localization microscopy *Opt. Express* **21** 10978–88
[21] Arganda-Carreras I, Sorzano C O S, Kybic J and Ortiz de Solorzano C 2008 bunwarpj: Consistent and elastic registration in imagej. methods and applications *ImageJ Conference Users-Developers*
[22] Schneider C, Rasband W and Eliceiri K 2012 NIH image to imagej: 25 years of image analysis *Nat. Methods* **9** 7
[23] Schindelin J *et al* 2012 Fiji: An open-source platform for biological-image analysis *Nat. Methods* **9** 676–82
[24] de la Rosa-Trevin J M *et al* 2016 Scipion: A software framework toward integration, reproducibility and validation in 3d electron microscopy *J. Struct. Biol.* **195** 93–9
[25] Jorge-Peñas A, Izquierdo-Alvarez A, Aguilar-Cuenca R, Vicente-Manzanares M, Garcia-Aznar J M, Van Oosterwyck H, de Juan-Pardo E M, Ortiz-de Solorzano C and Muñoz-Barrutia A 2015 Free form deformation-based image registration improves accuracy of traction force microscopy *PLoS One* **10** 12
[26] Maluenda D *et al* 2019 Flexible workflows for on-the-fly electron-microscopy single-particle image processing using scipion *Acta Crystallographica Section D: Structural Biology* **75** 882–94

Over the decades, biological image analysis, where microscopy innovation meets computer vision, has witnessed remarkable transformations, reshaping our understanding of life sciences. From the shift from traditional photography to digital imaging, laying the foundation for noise reduction, contrast enhancement, and object counting, to the establishment of fundamental algorithms in the 1970s for segmentation, object recognition, and feature extraction, the journey expands. A crucial moment emerged with artificial intelligence, enabling high-throughput analysis of vast microscopy datasets, ushering in a new era of discovery. The fusion of computer vision and biology gave birth to bioimage analysis, extracting quantitative insights from biological sample images. As data volumes surged, interdisciplinary collaborations led to specialized tools and platforms. This thesis focuses on advancing customization and automation in quantitative fluorescence microscopy image analysis at the Quantitative Image Analysis Unit (QIAU) belonging to the National Centre for Biotechnology (CNB-CSIC). This work contributes to the evolution of automated and quantitative optical microscopy analysis, offering potential for further integration with existing microscopy platforms, enhancing efficiency and user-friendliness in bioimage analysis.

Universidad
Complutense de
Madrid

www.ucm.es