



OPEN ACCESS

EDITED BY

Thomas Boudier,
Centrale Marseille, France

REVIEWED BY

Gabriele Gut,
University Hospital Zürich, Switzerland
Vannary Meas-Yedid,
Institut Pasteur, France
Juan Nunez-Iglesias,
Monash University, Australia

*CORRESPONDENCE

Kevin W. Eliceiri,
✉ eliceiri@wisc.edu

RECEIVED 20 May 2024

ACCEPTED 09 September 2024

PUBLISHED 27 September 2024

CITATION

Selzer GJ, Rueden CT, Hiner MC, Evans EL,
Kolb D, Wiedenmann M, Birkhold C,
Buchholz T-O, Helfrich S, Northan B, Walter A,
Schindelin J, Pietzsch T, Saalfeld S, Berthold MR
and Eliceiri KW (2024) SciJava Ops: an improved
algorithms framework for Fiji and beyond.
Front. Bioinform. 4:1435733.
doi: 10.3389/fbinf.2024.1435733

COPYRIGHT

© 2024 Selzer, Rueden, Hiner, Evans, Kolb,
Wiedenmann, Birkhold, Buchholz, Helfrich,
Northan, Walter, Schindelin, Pietzsch, Saalfeld,
Berthold and Eliceiri. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

SciJava Ops: an improved algorithms framework for Fiji and beyond

Gabriel J. Selzer¹, Curtis T. Rueden¹, Mark C. Hiner¹,
Edward L. Evans^{1,2}, David Kolb^{3,4}, Marcel Wiedenmann^{3,4},
Christian Birkhold^{3,4}, Tim-Oliver Buchholz^{5,6}, Stefan Helfrich⁴,
Brian Northan⁷, Alison Walter^{1,2,4}, Johannes Schindelin^{1,2,5,8},
Tobias Pietzsch⁵, Stephan Saalfeld^{5,9}, Michael R. Berthold^{3,4} and
Kevin W. Eliceiri^{1,2*}

¹Center for Quantitative Cell Imaging, University of Wisconsin–Madison, Madison, WI, United States, ²Morgridge Institute for Research, Madison, WI, United States, ³Department of Computer and Information Science, University of Konstanz, Konstanz, Germany, ⁴KNIME GmbH, Konstanz, Germany, ⁵Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany, ⁶Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland, ⁷True North Intelligent Algorithms, Guilderland, NY, United States, ⁸Microsoft Corporation, Redmond, WA, United States, ⁹Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA, United States

Decades of iteration on scientific imaging hardware and software has yielded an explosion in not only the size, complexity, and heterogeneity of image datasets but also in the tooling used to analyze this data. This wealth of image analysis tools, spanning different programming languages, frameworks, and data structures, is itself a problem for data analysts who must adapt to new technologies and integrate established routines to solve increasingly complex problems. While many “bridge” layers exist to unify pairs of popular tools, there exists a need for a general solution to unify new and existing toolkits. The SciJava Ops library presented here addresses this need through two novel principles. Algorithm implementations are declared as plugins called Ops, providing a uniform interface regardless of the toolkit they came from. Users express their needs declaratively to the Op environment, which can then find and adapt available Ops on demand. By using these principles instead of direct function calls, users can write streamlined workflows while avoiding the translation boilerplate of bridge layers. Developers can easily extend SciJava Ops to introduce new libraries and more efficient, specialized algorithm implementations, even immediately benefitting existing workflows. We provide several use cases showing both user and developer benefits, as well as benchmarking data to quantify the negligible impact on overall analysis performance. We have initially deployed SciJava Ops on the Fiji platform, however it would be suitable for integration with additional analysis platforms in the future.

KEYWORDS

algorithms, image analysis, scientific imaging, imaging software, analysis workflows, Fiji, OpenCV, python

1 Introduction

Scientific image analysis is, at its core, the act of applying algorithms to image data to explore a hypothesis. Image analysts need to extract knowledge from raw image data, the numerical signals recorded from an instrument. To do this, they need algorithms: mathematical routines to interrogate, transform, and distill the data into more refined forms; for example, segmenting a biological image into regions of interest (ROI) such as cells. To apply these algorithms at scale across large quantities of image data, they must be expressed as computer programs. Because building computational tools is a lengthy and painstaking process, software developers construct solutions on top of existing foundations whenever possible. But there are many diverging foundational layers from which to choose, including various machine architectures, operating systems, programming languages, and software libraries. Complicating matters, image datasets are stored in a variety of ways, such as on the cloud (Ouyang et al., 2023), on the GPU (Haase et al., 2020), or in memory; their values could be represented in numerous ways such as 8-bit integers (0–255) or floating point (a wide range but with some limits on numerical precision); and these values could be organized along rows, columns, strips, tiles, or in multidimensional blocks in one of many different file formats (Linkert et al., 2010; Hiner et al., 2016; Moore et al., 2021). Each choice imposes its own costs and benefits, and to allow computation within different settings a huge number of image analysis routines have naturally proliferated.

As the quantity and diversity of algorithm implementations has grown, there have been many efforts to organize them into cohesive software platforms to the benefit of users. Such platforms serve to unify access to compatible algorithms by operating on common data structures with defined conventions, making it easier to link them into workflows: sequences of operations where each algorithm is executed as a step in a larger analysis process. But each platform must be grounded in specific technical choices. Scientific communities tend to be partitioned by programming language as integration across language boundaries adds substantial complexity to software design. Some bioscience platforms build on the Java Virtual Machine (JVM), including ImageJ (Schneider et al., 2012), Fiji (Schindelin et al., 2012), QuPath (Bankhead et al., 2017), and Icy (De Chaumont et al., 2012). Many others build on the Python programming language and its n -dimensional NumPy array data structure (Harris et al., 2020) and SciPy extensions for scientific computing (Virtanen et al., 2020), including the scikit-image library (van der Walt et al., 2014), CellProfiler (Stirling et al., 2021), and napari (Ahlers et al., 2023). Web-based tools are also increasingly prevalent: examples in the bioimaging space include ImJoy (Ouyang et al., 2019), the BioImage Model Zoo (Ouyang et al., 2022), Cytomine (Rubens et al., 2019), and BIAFLOWS (Rubens et al., 2020). In the field of medical imaging, C++-based software platforms have been developed and refined for decades, including the Insight Toolkit (ITK) (Yoo et al., 2002) and Visualization Toolkit (VTK) (Schroeder et al., 2000).

Some platforms strive to integrate functionality across paradigms—e.g.: Jupyter Notebook (Kluyver et al., 2016) offers a visual browser-based application for developing and disseminating workflows written in a variety of languages including Python, Java, and others; KNIME Analytics Platform (Berthold et al., 2008) is an extensible tool for data science including image processing (Dietz

et al., 2016) which combines a JVM-based desktop application with web-based server components and support for executing Python programs as workflow nodes; and OMERO (Moore et al., 2015) is a web-based platform for managing microscopy image data, accessible from Java, Python, C++, and MATLAB programs as well as via its OMERO.insight desktop application. Even with such integrations in place, however, most platforms are not naturally mutually compatible: e.g., an “image” in ITK is not the same thing technically as an ImageJ image or a NumPy array. This heterogeneity of image data structures can be frustrating for users, who often struggle to combine multiple tools, leading to the creation of numerous additional integration mechanisms to link otherwise incompatible platforms (Dietz et al., 2020; Dobson et al., 2021; Hiner et al., 2017; Domínguez et al., 2017). Fortunately, Python has emerged as a common integration point for many scientific software platforms across these languages, with research software engineers putting in the extra effort required to make tools work from Python programs in addition to their original target platforms. The rise of deep-learning-based image analysis in the latter half of the 2010s (Li et al., 2023; Lambert and Waters, 2023; Nogare et al., 2023) further cemented Python as effective common ground to unify algorithm access. In the case of ImageJ and Fiji, for example, we created the PyImageJ library (Rueden et al., 2022) and napari-imagej plugin (Selzer et al., 2023) to facilitate easier access to ImageJ and other JVM-based routines from Python programs. ITK similarly has developed Python bindings (Lowekamp et al., 2013), as well as better integration into web technologies (Yaniv et al., 2018), to make its algorithms more accessible.

For users, who need to find the most efficient algorithm to accomplish their tasks, this wealth of tools is itself a barrier to task completion. The most efficient algorithm may be **inaccessible**, written in an incompatible programming language or library, or written for unavailable hardware. Algorithms can additionally be poorly documented or abandoned, making them hard to find and use even if they could be run. Finding the most efficient algorithm can also be challenged by **advancement**, as novel approaches and new hardware offer new capabilities. Developers naturally utilize new technologies, creating a continuous feed of new software components to consume, and an increasingly large collection of abandoned tools. This environment continually pressures users to learn new technologies, consuming valuable time and effort that could instead be directed at real applications.

The ImageJ Ops library, part of the ImageJ2 project aiming to expand ImageJ’s utility (Rueden et al., 2017), instead defines algorithms as structured plugins. In this way, it reframes image processing as a declarative task, where users decide which algorithms to apply and the ImageJ Ops infrastructure selects the best implementations for particular input data. While this approach made it very easy to contribute and grow libraries of algorithms, ImageJ Ops had several limits to its accessibility: its declarative matching could not take full advantage of type knowledge including Java generics or return type safety, plugin declaration through Java annotations presented a barrier to inclusion for other programming languages, and algorithm implementations were expected to build upon ImageJ’s data models. The SciJava Ops project described in this manuscript is a re-imagining of the ImageJ Ops design, as a declarative mechanism for retrieving and executing scientific algorithms. Figure 1 shows the place of Ops in the scientific

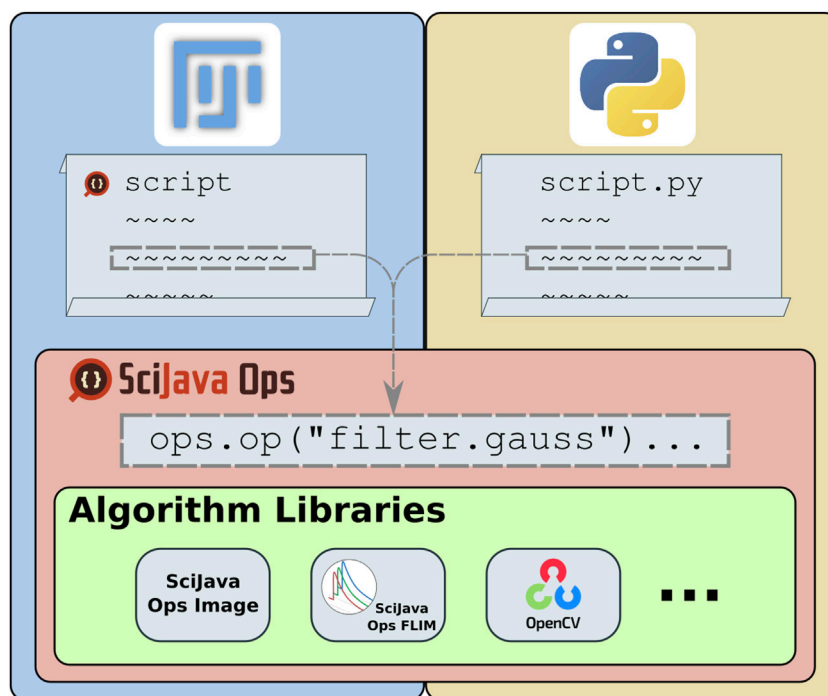


FIGURE 1
 SciJava Ops provides access to a diverse suite of algorithms through a uniform syntax which allows the framework to match the most appropriate algorithm to each user request and allows users to freely combine otherwise incompatible algorithms into a streamlined workflow. All SciJava Ops algorithms are usable from any of Fiji’s supported script languages (e.g., Groovy or Jython), as well as from Python scripts.

technology stack: providing a framework for algorithm discovery and application, agnostic of data model or programming language, creating a unifying point of access to algorithm libraries.

2 Method

Like ImageJ Ops, SciJava Ops aims to abstract the task (e.g., a Gaussian blur of an image with a given sigma value) from the concrete mechanisms (using scikit-image with zarr arrays) by providing a set of framework features. This declarative approach allows both advancement, as novel algorithms can replace existing algorithms without affecting the abstraction, and also accessibility, when algorithms across languages, libraries and data structures can interact through a unified framework. To provide a viable declarative approach, however, SciJava Ops must define framework features and structure that can encompass the syntax and semantics of algorithms across a diverse set of mechanisms, allowing users to interact with all algorithms in a unified way and additionally allowing unified interactions between the algorithms themselves. This section describes the fundamental abstractions and structure that comprise the SciJava Ops framework.

2.1 Design principles

SciJava Ops is built atop the following abstractions, designed to encompass algorithms across language, library, and hardware:

1. An **Op** is a plugin defining an algorithm implementation. It has a **name** indicating its purpose, and a variable number of input and output **parameters**, each defined by a language-specific **type**. For example, a function that adds two integers, *a* and *b*, producing the result as a new integer, might be represented by the Op `math.add(a:int, b:int) → int`. Ops adhere to **functional interfaces**: structural patterns that define common behavior.
 - A **function Op** accepts immutable (unchanging) **input** values and produces a new **output** value. E.g.: `filter.gauss(input:image, sigma:float) → image` accepts an existing image as input, along with a real-valued sigma, and performs a Gaussian blur operation, producing a transformed image as new data, occupying additional computer memory, while leaving the original input image unchanged.
 - A **computer Op** accepts immutable input values and writes its output into a pre-allocated **container**. E.g.: `filter.gauss(input:image, sigma:float, output:image)` accepts an existing image as input, along with a real-valued sigma, and performs a Gaussian blur operation, overwriting the contents of the container image (named `output`) with the transformed data. In this way, no new computer memory needs to be allocated for the computation.
 - An **inplace Op** accepts input values, one of which is **mutable** and overwritten with the transformed data. E.g.: `filter.gauss(input:image, sigma:float)` performs a Gaussian blur operation on the input image,

overwriting its contents with the computation result, thus destroying the original data values, but saving space by avoiding allocation of additional computer memory.

- Each of these three patterns have advantages and disadvantages, and as each is frequently seen (though typically less formally) in existing libraries of algorithms, all three are necessary to accommodate the broadest range of potential Ops.
- Ops are declared via metadata descriptor files, enabling all Ops to be cataloged in a unified way, independent of language and implementation. Each entry in one of these descriptor files references an Op written for a particular target platform, such as a Python function or Java method. While these files alone are not enough to actually execute any particular Op, they are sufficient to reason about which Ops would be available. Descriptor files also enable **zero-code wrapping**: algorithmic code from existing libraries can be made available within SciJava Ops *without being written for SciJava Ops*.
- All available Ops are collected into an Op **environment** which provides access to Ops functionality. Via this environment, Ops are **requested** by name and parameter types. The intent of the design is to allow users to specify the algorithm they want to run, without mandating which specific implementation will be used. This approach leaves room for the best implementation to be chosen each time an algorithm is requested, enabling existing Ops-based workflows to automatically improve over time as better Ops are added to the environment. For example, an image processing script that calls `filter.gauss` may become faster when a GPU-based Gaussian filter Op becomes available.
- Op execution requests are powered by the Ops **engine**: a matching system that chooses the most appropriate available Op from the environment. The engine takes into account not only the number and types of parameters, but also the type of functional interface (function, computer, or inplace). In case no Op exists with strictly matching parameter types, the engine can implicitly convert parameters into other types that *do* match an existing Op. For example, an Op written to convolve an ITK image could be used even when the user requests to convolve an image opened by Fiji, without the user needing to be cognizant of ITK data structures, as long as the environment has `engine.convert` Ops available capable of converting between those two types of images. The Op matching system is **deterministic**, necessary for **reproducible** science; given a fixed collection of Ops within the Op environment, the same execution request will always be matched to the same Op.
- Ops are atomic units that can build on each other. For example, a Difference of Gaussians (dog) Op:

```
filter.dog (input: image, signal: float,
           sigma2: float) → image
```

might leverage helper Ops as **Op dependencies** to execute the formula:

```
filter.gauss (image signal) - filter.gauss
(image, sigma2)
```

Where the minus (-) operator is executed using a `math.sub (a, b)` Op, as shown in Figure 2C below. Such dependencies are also decided by the engine, rather than linking to specific Ops. For example, if the more performant GPU-based `filter.gauss` Op described in Linkert et al. (2010) above is added to the environment, then the `filter.dog` Op described here would also benefit automatically.

2.2 Architecture

The Ops design elements above are independent of any programming language—but ultimately the design must be implemented as a software library built on a concrete foundation. We chose the Java Virtual Machine (JVM) to implement the Ops design for several reasons: 1) the widespread adoption and ease of use of the Fiji platform, written in Java, makes it an ideal initial deployment target across computing platforms (Schindelin et al., 2015); 2) the ImageJ Ops library served as a launching point to improve on the existing Ops design while providing a large number of image processing Ops built on the powerful n-dimensional ImgLib2 library (Pietzsch et al., 2012) as a foundation; and 3) ImageJ and Fiji functions are already accessible from Python via the PyImageJ project, offering an immediate way for Ops to be used from Python as well, and later to wrap existing Python libraries into Ops. Nonetheless, this implementation could serve as a viable blueprint for non-JVM environments in the future, should the need arise.

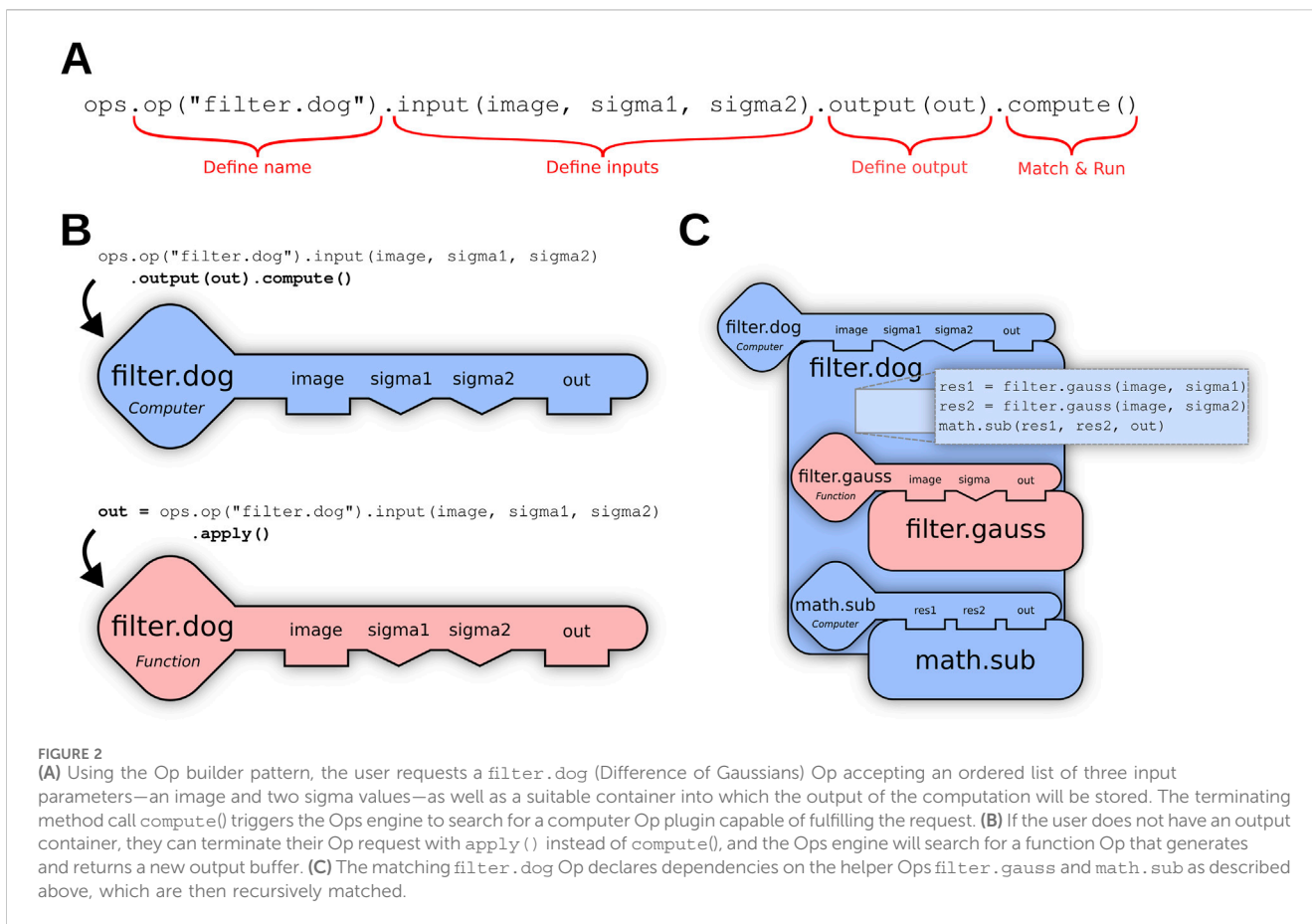
SciJava Ops is partitioned into several different component libraries. The following sections describe the application programming interface (API) layers of SciJava Ops in increasing technical concern. We start with basic querying and use of Ops, move to development of new Ops, and end with technical overviews of the mechanisms powering Ops from user and developer perspectives.

2.2.1 Requesting Ops

The most common entry point into SciJava Ops is the `scijava-ops-api` component, which offers a way to request Ops declaratively by name and parameter types, using the **Op builder** mechanism of the Op environment class. The Op builder pattern partitions the request into a granular set of function calls, as illustrated in Figure 2A; completed Op builder requests are then provided to internal SciJava Ops components to attempt to match an algorithm implementation.

Several variations of the above pattern are possible, depending on whether the user wants to match and execute the Op immediately, or receive a reference to the matched Op that can be invoked repeatedly as desired. For example, if the query is terminated with `.function()` rather than `.apply()`, the matching `math.add` Op itself will be returned as a `BiFunction` object, Java's functional interface for binary functions. Other variations, as shown in Figure 2B, may be relevant when the user wants to match as an inplace or computer Op, rather than a function Op.

To assist the user with knowing which Ops are available to call, the Op environment also provides `.help()` and `.helpVerbose()` methods to obtain descriptions of all Ops



matching a particular query. For example, `ops.help("math")` would list all Ops in the math category (e.g., `math.add` and `math.sub`). On the other hand, `ops.help("math.add")` would list all known implementations of the `add` Op in the math category, while `ops.helpVerbose("math.add")` would provide the same list but with detailed usage documentation. These help methods can be called at any stage of the builder, listing potential Op matches according to current criteria.

One challenge here is that while all image analysis software ultimately operates on images, the naming of image data structures between software packages is inconsistent; for example, images are represented using `ImgLib2 RandomAccessibleInterval` structures within Fiji, as `ndarray` objects within NumPy-based projects, and as `Mat` objects within OpenCV (Bradski, 2000). Transparent integration and interoperability between these platforms requires not only strategies for conversion between these data structures, as described above, but also abstraction from platform-specific nomenclature. SciJava Ops addresses this issue via a **type description** mechanism fostering consistency of naming across data types, with the set of `engine.describeOps` map each type to a common, simplified name; for images, we might use `engine.describeOps` to describe `RandomAccessibleInterval`, `ndarray`, and `Mat` all as "image". Then, when the user asks for information about a given Op via `ops.help` or `ops.helpVerbose`, each

parameter is replaced with its description, resulting in text without any terminology specific to particular software packages.

2.2.2 Defining Ops

A critical design goal of SciJava Ops is to make it as convenient as possible to bring new Ops into the system. To extend the collection of Ops available for use, developers can implement new algorithms, as well as introduce existing libraries into the framework, in ways that require minimal to zero additional changes to source code or project dependencies.

For developers wishing to declare their routines as Ops directly in their implementation code, the `scijava-ops-indexer` component provides a simple way to do so using Javadoc: a special comment describing the purpose and behavior of a piece of Java code. A Java class, method, or field tagged with an `@implNote` line will trigger automatic production of an Ops metadata descriptor file in YAML format, thus making the Ops engine aware of the routines and their structure. Figure 3 depicts a static method that uses standard syntax and data structures, but includes an `@implNote` line in its Javadoc which registers the method as a `copy.array` computer Op.

Additional arguments, including the Op's name and priority, can be appended to this tag. It is important to note that this is a lightweight mechanism, allowing libraries to declare algorithms as Ops without explicitly requiring the Ops framework to be present to build or run their code outside of the Ops paradigm.

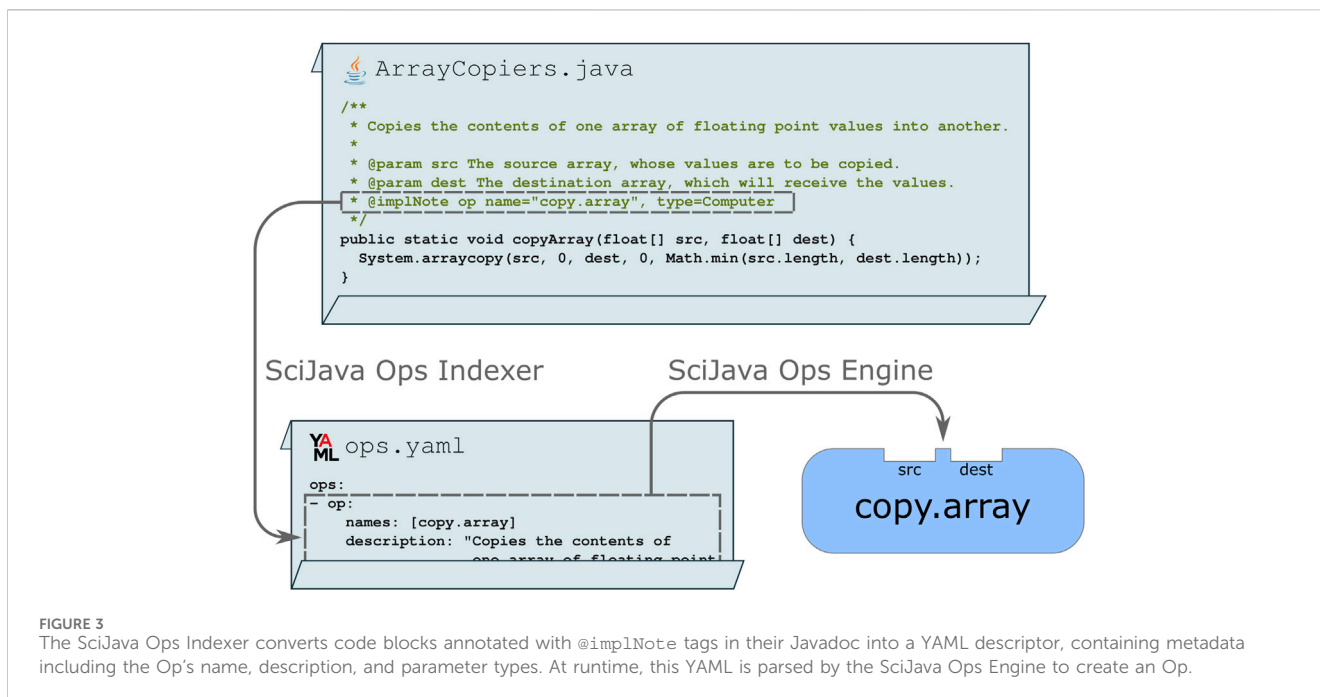


FIGURE 3 The SciJava Ops Indexer converts code blocks annotated with @implNote tags in their Javadoc into a YAML descriptor, containing metadata including the Op’s name, description, and parameter types. At runtime, this YAML is parsed by the SciJava Ops Engine to create an Op.

The Javadoc-based mechanism for defining Ops is convenient, but only usable by algorithm libraries written in Java, and requires modification to the source code’s Javadoc sections. More generally—and especially in cases where such modifications are undesirable or impossible—the needed Ops YAML metadata descriptor files can be produced in other ways, making the Ops framework aware of code that was not originally written with Ops in mind. For example, third party libraries such as OpenCV can be made accessible as Ops without modifying the OpenCV code (see “OpenCV” in the “Example Use Cases” section below).

2.2.3 Matching Ops

The Ops engine includes several layers known as **matching routines**, which provide powerful and flexible fulfillment of Op requests in ways beyond only the base implementing algorithm. A request to add two images element-wise can be fulfilled by an Op that adds two elements, which is “lifted” or mapped to the entire image. A request to convolve MATLAB matrices can be fulfilled by an Op that convolves ImgLib2 images, as long as the engine has access to engine.convert Ops that translate between those two image types. A request for a function Op to perform a Gaussian filter on an image can be fulfilled by a computer Op as long as there is an engine.create Op to produce the computer’s pre-allocated container output. In the following sections we will give more detail on these and related mechanisms.

2.2.3.1 Direct Op matching

This matching routine seeks an Op that can run directly on user inputs. Searching through all Ops with a name matching the user request, the default routine asserts that the input argument types conform to the required parameter types of the Op. While allowing declarative access, direct matching does not perform any more sophisticated transformations to the arguments or to the underlying Op.

2.2.3.2 Op adaptation

When no direct match exists for a user request (e.g., Figure 4A), further matching routines transform available Ops to produce a match: for example, calling an Op as if it implemented a different functional type. We define this as **Op adaptation**, enabled by special engine.adapt Ops to perform conversions as needed. Using adaptation, users can indirectly execute a computer Op as a function Op or invoke element-wise Ops across an entire list or image, embedding these transformations system-wide instead of requiring modifications per algorithm. Figure 4B visually depicts how Op adaptation enables new matching possibilities; practical examples can be seen in Section 4.2, where Op adaptation is used to execute pixel-wise and neighborhood-wise Ops across entire images as well as to create pre-allocated output buffers.

2.2.3.3 Parameter conversion

While Op adaptation alters the Op structure, a third matching routine called **parameter conversion** facilitates the conversion of individual inputs and outputs, resolving the situation shown in Figure 5A when a single parameter prevents an Op match. The Ops engine’s matching mechanism is fundamentally built around knowing the parameter types of each Op; with this knowledge, multiple ops of the same name can exist that operate on different data types, such that the best Op for the job can be chosen in different scenarios. But with parameter conversion we can also take it a step further: rather than restricting Op requests to types aligned with each implementation—requiring users to exercise prior knowledge of the implementation details of available Ops—the engine can convert input arguments on the fly between types, such that a broader range of available Ops can be matched than would otherwise be possible. For example, the logic of convolving an ImgLib2 image with an OpenCV image is not truly different from convolving two ImgLib2 images: the application of a single ImgLib2 convolve Op on images of either or both image

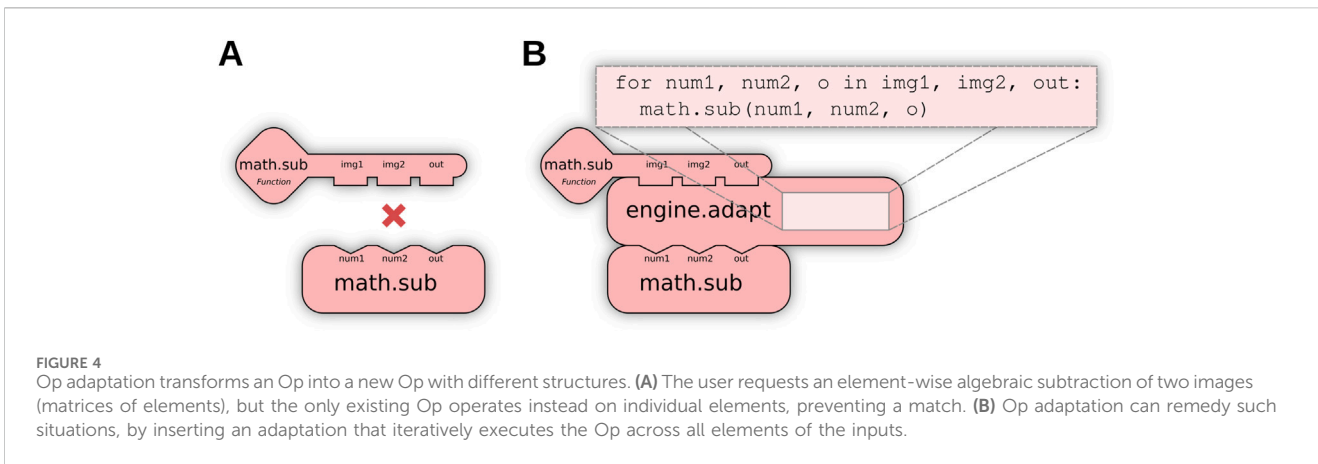


FIGURE 4 Op adaptation transforms an Op into a new Op with different structures. (A) The user requests an element-wise algebraic subtraction of two images (matrices of elements), but the only existing Op operates instead on individual elements, preventing a match. (B) Op adaptation can remedy such situations, by inserting an adaptation that iteratively executes the Op across all elements of the inputs.

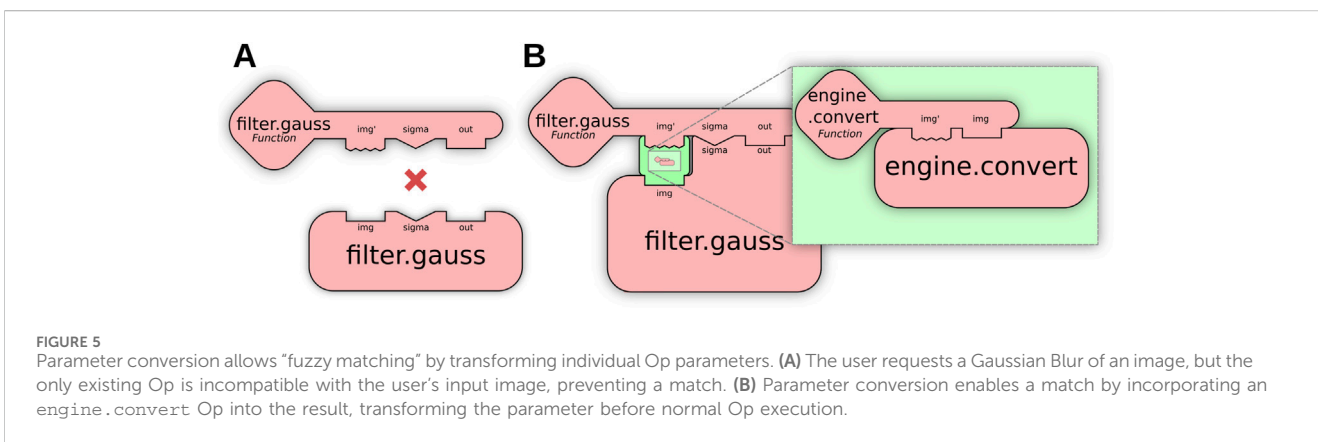


FIGURE 5 Parameter conversion allows “fuzzy matching” by transforming individual Op parameters. (A) The user requests a Gaussian Blur of an image, but the only existing Op is incompatible with the user’s input image, preventing a match. (B) Parameter conversion enables a match by incorporating an engine.convert Op into the result, transforming the parameter before normal Op execution.

types represents the core SciJava Ops principle of interoperability.

Parameter conversion utilizes engine.convert Ops to convert a user’s arguments to the Op implementation’s required types, and similarly converts Op outputs to the user’s requested output type. Figure 5B shows how these Ops can be inserted by the matcher to enable “fuzzy” matches, allowing users to execute Ops on a wider range of parameter types, removing the need for manual conversion beforehand or knowledge of exact Op implementations. A practical example of parameter conversion is described in Section 4.1, where it allows seamless execution of OpenCV algorithms using ImgLib2 data structures.

2.2.3.4 Optional parameters

Optional function parameters are a convenient feature of many popular programming languages. For example, a rescale2D(image, width, height) function for resizing a 2-dimensional image plane might mark the height parameter as optional, so that rescale2D can be invoked with only image and width arguments, calculating the target height in a way that maintains the image’s aspect ratio. SciJava Ops allows the use of optional parameters in an analogous fashion by creating “reduced” versions of the Op that omit parameters marked as optional one by one from right to left, in line with how most programming languages support

this feature. For instance, one of the Fast Fourier Transform (FFT) Ops in SciJava Ops Image has the form filter.fft(input, fftType, borderSize, fast) → output, where the borderSize and fast parameters are both optional and each making a reasonable assumption if left off; this results in additional “reduced” Ops filter.fft(input, fftType, borderSize) → output and filter.fft(input, fftType) → output also being available in the Op environment. Optional parameters enable Ops to provide reasonable default behaviors while matching against more concise Op requests.

2.2.4 SciJava Ops image

The scijava-ops-image component provides a large collection of image processing algorithms built on SciJava Ops, including functionality such as image filtering, thresholding, binary and grayscale image morphology operators, image segmentation and labeling, 2D and 3D object geometry, shape descriptors and measurements, numerical and image arithmetic, fast Fourier transformations, image convolution and deconvolution, spatial transformations, and colocalization. This library, which represents an evolution of the ImageJ Ops collection, is not intended to be exhaustive but instead foundational, offering Ops built on the ImgLib2 image library; other Ops built on other image libraries are equally possible.

3 Results

All components of the SciJava Ops framework are implemented in Java and deployed as libraries to the Maven Central repository, making them freely available for other software to build upon. SciJava Ops is also available in Fiji via the SciJava Ops update site, which provides access to and integration with its large and diverse plugin ecosystem (Cardona et al., 2012; Arzt et al., 2022; Hörl et al., 2019; Arena et al., 2017; Ershov et al., 2022; Arshadi et al., 2021; Legland et al., 2016; Domander et al., 2021). Detailed installation and usage instructions, as well as how-to guides and examples, can be found on the project website at <https://ops.scijava.org/>.

Through the architecture described above, SciJava Ops provides a framework for declarative, extensible, reusable algorithms, with a focus on image analysis. The following use cases describe how both users and developers benefit from SciJava Ops. The uniform syntax of SciJava Ops provides users the freedom to run algorithms from various sources, as Op calls are identical across library (Section 4.1) and programming language (Section 4.2). At the same time, developers obtain accessibility, as they can expose complete libraries within SciJava Ops. These libraries do not need to adhere to the SciJava Ops framework (Section 4.1), but can utilize powerful SciJava Ops features when they do (Section 4.3). Out of the box, SciJava Ops already provides powerful algorithms able to derive new meaning from modern data (Sections 4.4, 4.5).

3.1 Wrapping an existing image processing library to Ops: OpenCV

The `scijava-ops-image` component provides a foundation for image analysis, but much like ImageJ and Fiji, the strength of SciJava Ops lies in its potential for extensibility. However, where ImageJ has provided a template for extension within its ecosystem, the vision for Ops is a grand unification of image processing libraries with integration layers built on top as needed, instead of requiring adherence to a shared implementation contract. In practice this has been accomplished through the use of YAML to identify Ops algorithms: the YAML serves as a separate deliverable which can be defined independently of the implementation library it references; as long as both are accessible to the SciJava Ops framework, the defined Ops will be available for use. To provide a practical example of this potential we developed the `scijava-ops-opencv` component.

OpenCV is a massive open source library for computer vision. Although OpenCV is written in C++, it has bindings for other languages—we targeted the `bytedeco/javacv` (<https://github.com/bytedeco/javacv>) Java wrappers for convenience. The bare minimum for inclusion into SciJava Ops was a Maven dependency on the `bytedeco` libraries and the YAML configuration file mapping the methods therein to Ops. There are over 2,500 algorithms defined in OpenCV so we chose to focus our initial mappings to the Image Processing and Photo domains. Creating the YAML configuration file could be done manually, but this would be labor-intensive and tedious. Instead, we created a simple shorthand parsing program allowing the specification of OpenCV classes, methods, and essential metadata (e.g., the

functional type) which then automatically identified cases such as overloaded methods and output appropriate Op YAML definitions. Thus with a cursory look at the OpenCV codebase, we were able to expose several hundred new methods into SciJava Ops.

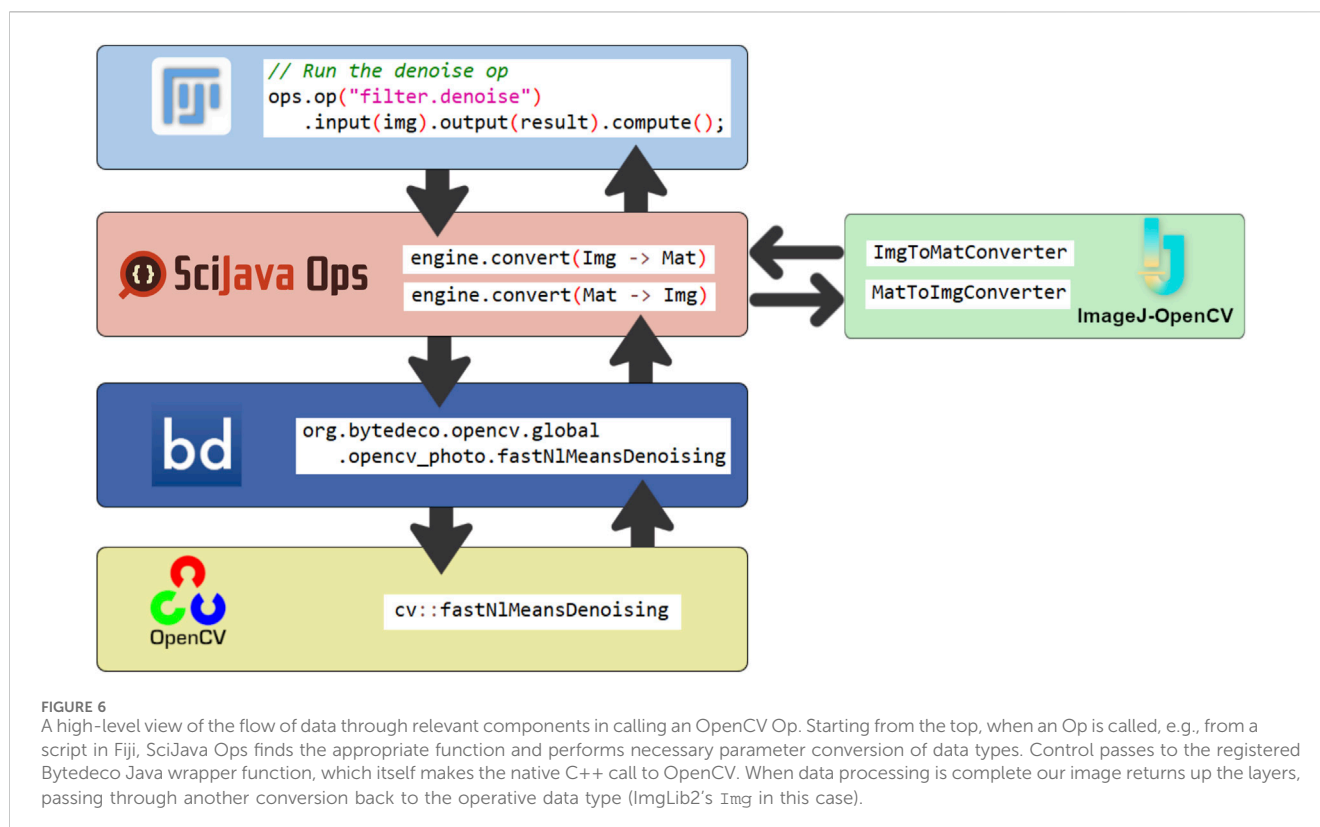
However, true integration and interoperability also requires an exchange of data. To that end, we leveraged the prior development effort of ImageJ-OpenCV (<https://github.com/imagej/imagej-opencv>) which defines translation between the Mat data structure of OpenCV and `RandomAccessibleIntervals` of `ImgLib2`. Similar reuse of our shorthand configuration allowed the inclusion of ImageJ-OpenCV's translation methods into the SciJava Ops framework. One final element was required by SciJava Ops to take advantage of its adaptation and parameter conversion mechanisms: the knowledge of how to copy between Mat data structures. Mats already have `copyTo` methods available in their API, so this simply required defining minimal copy Ops of our own that call this method. With these components in place, we could now use Mats and `RandomAccessibleIntervals` (and extensions thereof) interchangeably. Figure 6 provides a sense of how these software pieces all fit together when actually using an OpenCV Op.

There are several key observations highlighted in Figure 6. First, the OpenCV Non-Local Means denoise algorithm is being called in a general way through Ops: `filter.denoise` with the appropriate parameters. It is not identifiably tied to OpenCV. Second, we are providing the image data to the denoise algorithm as an `ImgLib2` data structure despite the fact that the underlying OpenCV implementation operates on native `cv:Mat` instances. This conversion is happening automatically, transparent to the calling user. Perhaps most importantly, we can see the keystone role of SciJava Ops in unifying disparate software efforts for the benefit of connecting researchers with increased algorithm diversity. This implementation is representative of the general requirements for including new libraries into SciJava Ops: creating converters between common data structures, and generating a YAML registering the library's methods as Ops.

3.2 Invoking sciJava Ops from Python: scjjava

While we have made SciJava Ops accessible within Fiji via an ImageJ update site, there is nothing ImageJ- nor Fiji-specific about its implementation. Indeed the vision for SciJava Ops has always been to unite algorithm libraries across analysis tools and programming languages. Python integration is of particular interest due to its prevalence in the scientific community and forefront position in the advancement of deep learning. Although the core of SciJava Ops is, naturally, written in Java, we can take advantage of the low-level `scjjava` and `jgo` libraries—built to facilitate `PyImageJ`—to gain immediate access from a Python entry point. Notably, SciJava Ops usage within a Python environment is identical to usage upon the JVM, allowing users to exercise familiarity with SciJava Ops within new workflow environments; once a user understands how to use SciJava Ops in one environment, they will be able to execute Ops in every environment.

In Figure 7 we see this process in-depth, starting from a minimal Python environment. As all SciJava Ops libraries are deployed to a public



Maven repository, the `scjjava` configuration mechanism can succinctly be instructed to gather all necessary Java code and make it available in Python. At that point we can freely call Ops (or other Java code) as needed. For the sake of example we show Ops in use to perform basic image preprocessing, segmentation and measurement. While these particular Ops were written to operate on `ImgLib2` data structures, the logic for converting from Pythonic data types (numpy) is already present in `PyImageJ`; registration of those methods as Ops would allow for seamless conversion and inter-use. Note that this use of `scjjava` would be appropriate from an interpreter or running as a script, but similarly enables SciJava Ops to be called from other Python tools, e.g., `napari`, as was done for the `napari-imagej` plugin.

3.3 Extending SciJava Ops with new libraries: fluorescence lifetime image analysis

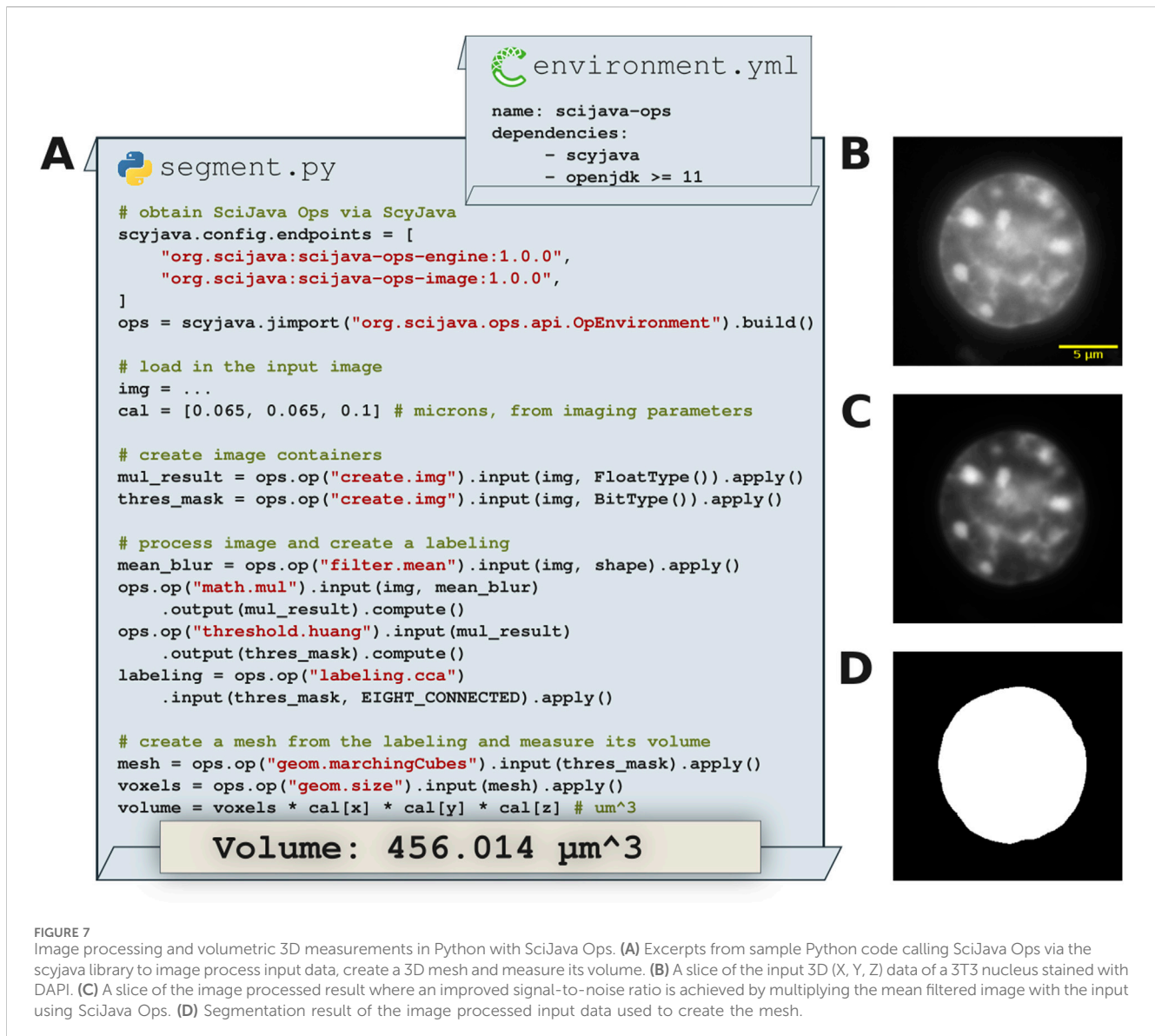
Fluorescence Lifetime imaging microscopy (FLIM) utilizes the excited state decay rate of fluorophores to derive additional information about the surrounding microenvironment. Open, accessible tooling for FLIM analysis in Fiji is provided by `FLIMJ` (Gao et al., 2020), an `ImageJ` plugin wrapping the C library `FLIMLib` for use with `ImgLib2` images. Included within `FLIMJ` is a set of algorithms which we have exposed within the SciJava Ops framework. The resulting Ops, contained within the supplementary SciJava Ops `FLIM` component, provide fluorescence decay rate fitting to one of many exponential-decay models, pseudo coloring of fit results, and subsequent post-processing. Through SciJava Ops's discovery

mechanisms, the resulting Ops appear as new entries in the Op environment, the same as any other Op.

Figure 8A depicts a SciJava script that performs a Levenberg-Marquardt algorithm (LMA) fit (TCSPC, 2021) and pseudo-coloring the fluorescence lifetime map, illustrating the transparent interoperability of SciJava Ops and existing Fiji components. Using the `Bio-Formats` plugin for Fiji (Moore et al., 2015), we load the image data into Fiji along with metadata required for FLIM analysis. Figure 8B shows how `ImageJ`'s selection tools enable drawing of any number of `ImageJ` ROIs, which are then converted using Fiji to the `ImgLib2` data structures used by SciJava Ops `FLIM`. In this case, the flexibility provided by these tools allows users to restrict computation to subregions of the data, as shown in Figure 8C, avoiding the penalties of computation across undesired or uninteresting areas. With this script, users can observe fluorescence and the fluorescence lifetime across the region of interest, compute parameters of interest and could additionally perform segmentation, labeling or other post-processing routines.

3.4 Making novel image analysis algorithms more accessible: spatially adaptive colocalization analysis

Colocalization analysis remains a powerful tool to study the co-occurrence of two signals of interest in a multi-channel dataset. Typically, workflows select a ROI, chart pixel data from both channels within the ROI on a scatter plot and determine the degree of colocalization by determining the colocalization quantification index (e.g., Pearson's, Li, etc.) (Manders et al., 1993; Li et al., 2004). While



effective, this strategy fails to account for the valuable spatial information in the rest of the image and relies upon a non-reproducible ROI. Wang et al. (2019) address these shortcomings through the Spatially Adaptive Colocalization Analysis (SACA) algorithm. Unlike traditional colocalization approaches requiring a predetermined ROI the SACA algorithm determines colocalization strength at the pixel level (Figure 9C), considering all pixels within the image. By doing so SACA takes advantage of the spatial information within the image, avoiding ROI selection entirely.

SACA was originally implemented in R, limiting its accessibility from other popular image analysis platforms in other languages such as Python and Java. To improve the situation, we translated the SACA algorithm into two Ops within the SciJava Ops Image library. The `coloc.saca.heatmapZScore` Op (Figure 9C) accepts two input images and returns the z-score heatmap of colocalization strength. Through incorporation within the Ops framework, the `coloc.saca.heatmapZScore` Op is able to depend upon `threshold.otsu` and `image.histogram` Ops, avoiding significant code duplication. The `stats.pnorm` Op accepts can be

used with the z-score heatmap as an input to create a pixel wise *p*-value heatmap (Figure 9D). Similarly the `coloc.saca.sigMask` Ops works on the same z-score heatmap output from SACA and returns a binary mask indicating which pixels are significantly colocalized (Figure 9E). Integrating the SACA algorithm with automated regression tests into the `scijava-ops-image` library not only makes it accessible from the Fiji platform and linked tools, but also helps foster continued maintenance and support over time.

3.5 Improving access to established image processing algorithms: Richardson-Lucy TV deconvolution

Although epifluorescence microscopy allows researchers to acquire sophisticated datasets, the images acquired still suffer from out-of-focus blur due to the light scatter through the sample. This problem can also be described as convolving the input image with the point spread function (PSF), the function

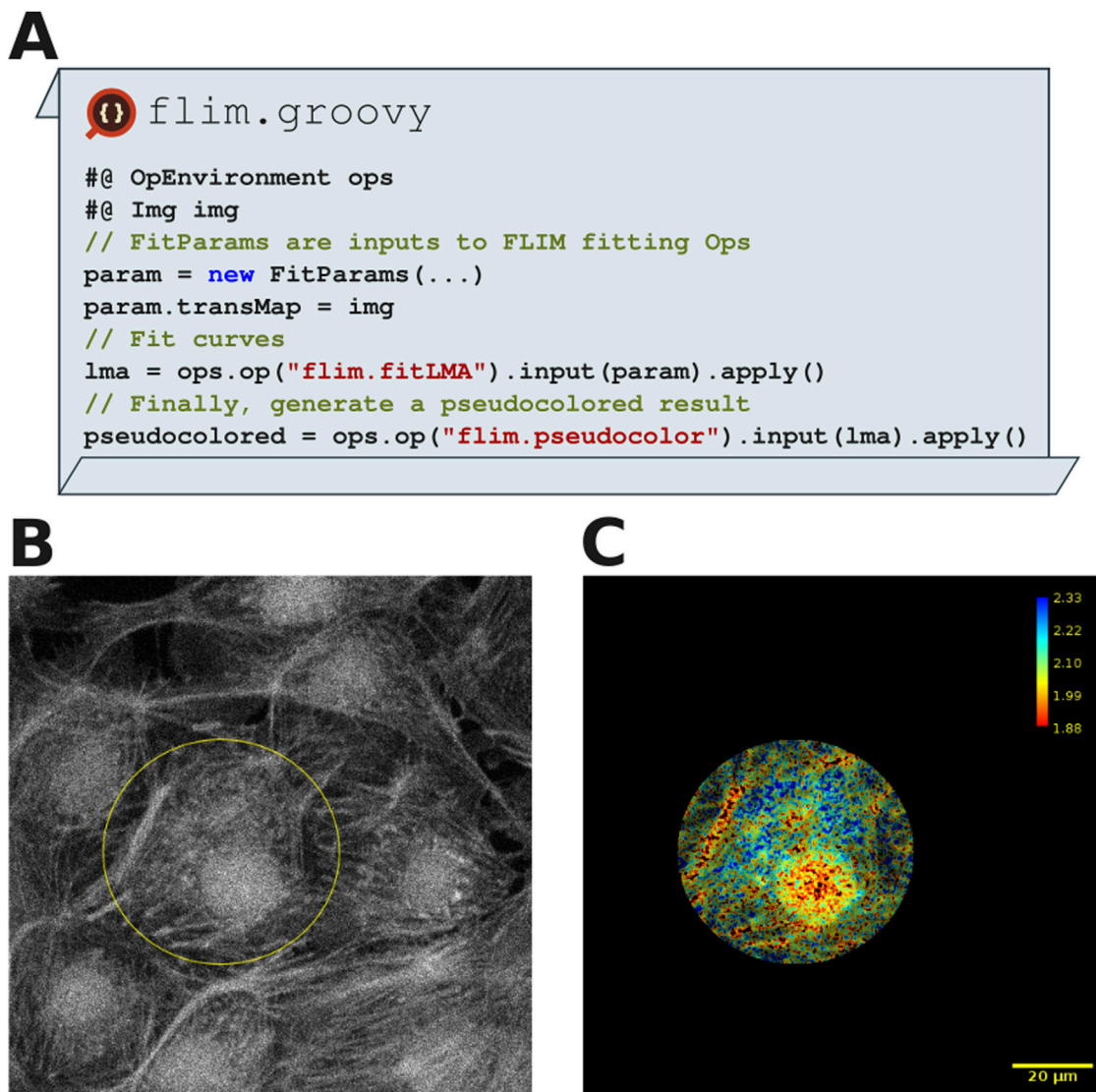
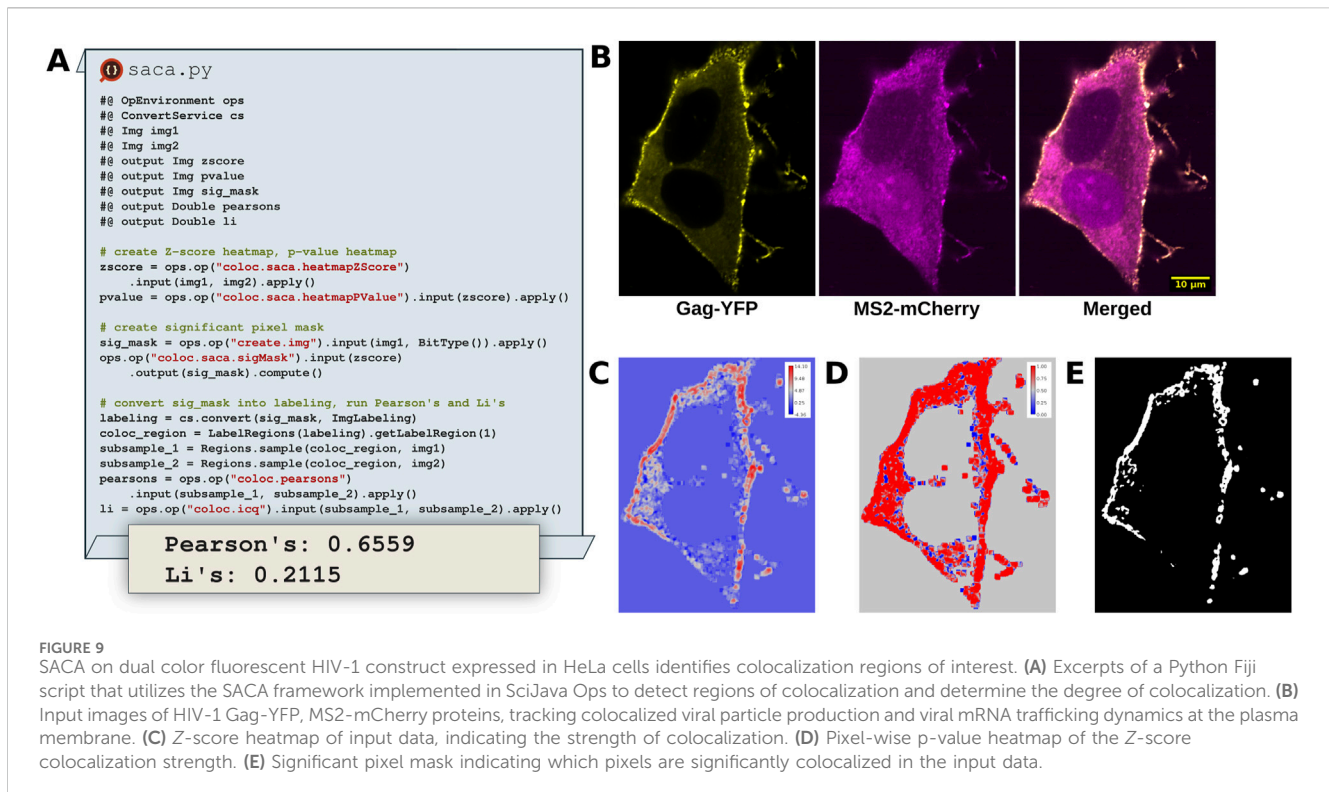


FIGURE 8
 FLIM analysis shows a separation of three fluorescence components in a MitoTracker Red, Alexa Fluor 488 and DAPI labeled fixed BPAAE cell sample (FluoCells Prepared Slide#1, ThermoFisher). **(A)** Excerpts from FLIM analysis script using SciJava Ops which was run in Fiji. **(B)** A 2D FLIM image of BPAAE cells, imported into Fiji using Bio-Formats. An elliptical ROI has been drawn around a single cell using the ImageJ Elliptical selection tool. **(C)** FLIM analysis within SciJava Ops yields pseudo colored result images (with corresponding color bar), which could be further processed using other Ops. Note that script results were restricted to the elliptical ROI drawn in **(A)**. The scale bar is 20 μm and the FLIM lookup table reads 1.88 ns–2.33 ns.

that describes how light diffracts through the optical system. By performing the inverse of convolution (*i.e.*, deconvolution) with the PSF, the out-of-focus blur can be removed, recovering the original image. Several plugins have been written for the original ImageJ platform providing different kinds of deconvolution, including Bob Dougherty's Diffraction PSF 3D and Iterative Deconvolution 3D; the EPFL's BioImaging Group's PSF Generator, DeconvolutionLab, and DeconvolutionLab2; and Piotr Wendykier's Parallel Iterative Deconvolution and Parallel Spectral Deconvolution (Wiki, 2024). Unfortunately, none of them implement the gold standard deconvolution algorithm, Richardson-Lucy TV (Dey et al., 2006), nor the gold standard point-spread function generation algorithm, Gibson-Lanni (Gibson and Lanni, 1991).

The SciJava Ops Image module implements both Richardson-Lucy (RL) and Richardson-Lucy Total Variation (RLTV) deconvolution (Dey et al., 2006), which improve image restoration quality by employing a regularization factor to suppress noise normally amplified by the RL algorithm (compare Figures 10B, C). Both implementations optionally allow vector acceleration (Biggs and Andrews, 1997) to speed up convergence and non-circulant boundary handling (Bertero and Boccacci, 2005) to reduce artifacts near edges. SciJava Ops Image also implements the Gibson-Lanni algorithm, `createKernelDiffraction`, to create a synthetic PSF (Figure 10D) with given imaging parameters including wavelength, and numerical aperture. Taken together, SciJava Ops allows researchers to perform 2D and 3D deconvolution with one framework.



4 Discussion

4.1 Scalability

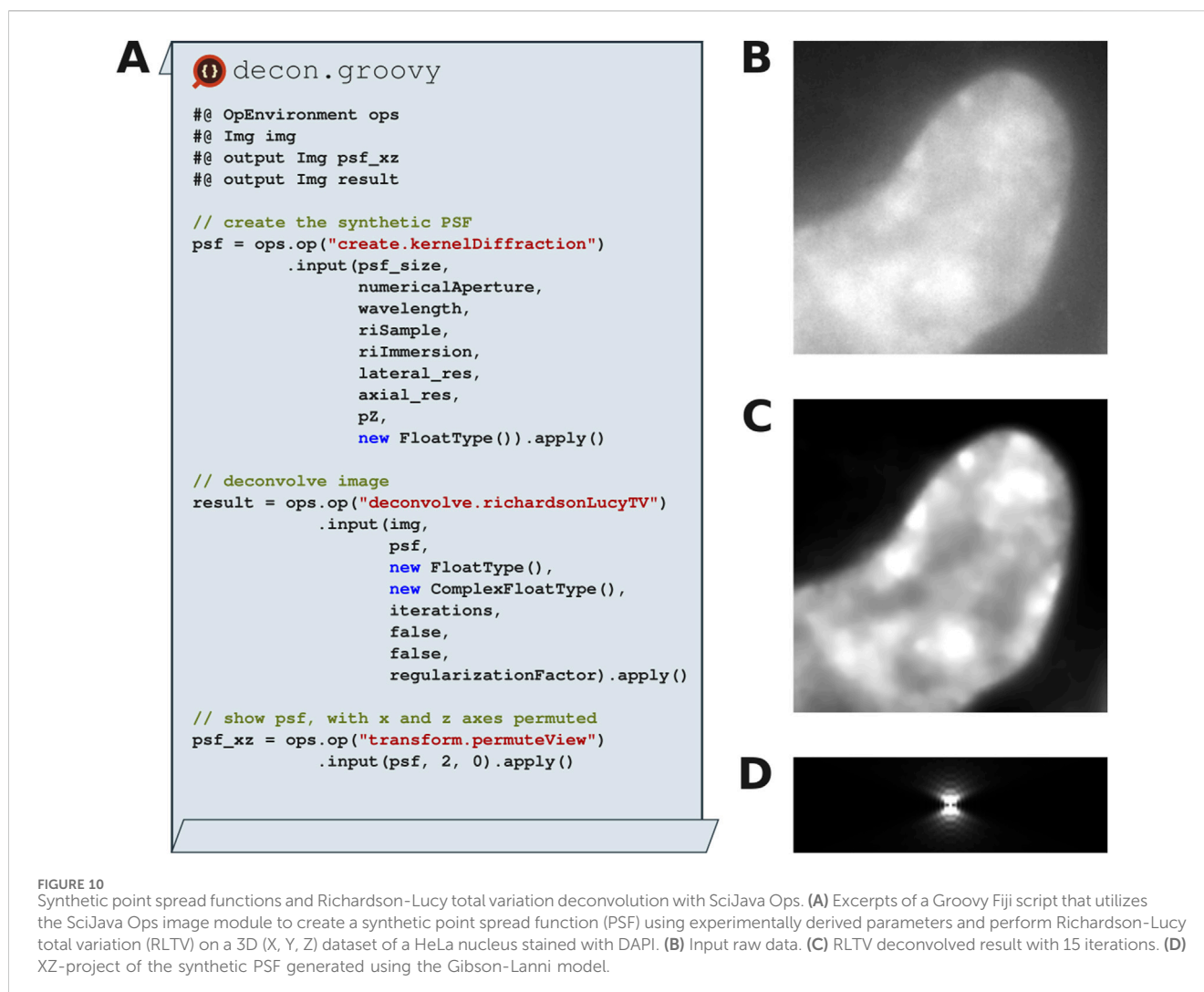
When paired with the ImgLib2-cache framework, SciJava Ops can be used to generate or process images of virtually unlimited size. ImgLib2 supports very large, tiled images (`LazyCellImg`) where tiles (cells) are lazy-loaded when they are first accessed. A typical usage scenario is to load cells from disk on demand and keep recently used cells in a memory cache. However, cells can also be generated by arbitrary computations, cells can be modified and modifications written to disk when cells are evicted from cache, etc. Importantly, this all happens transparently, internal to ImgLib2, and is presented as a normal ImgLib2 image (`RandomAccessibleInterval`). This interacts with SciJava Ops in several ways:

First, such cached images can be used as input and output arguments for any Op on images. Many Ops can then simply work on larger-than-memory images, in particular Ops that use ImgLib2 data structures. Second, cells can be generated by Ops. Consider, for example, applying a per-pixel intensity correction to an input image. The result of this operation can be a `LazyCellImg` (with the actual computation only happening when pixels of this result are accessed). The cell loading functions would then run the intensity correction Op with the cell as the output image. Note, that in this case even Ops that require argument materialization can be applied, because they are run only on small portions of input/output image at a time. Finally, because the result of such an operation is just a (lazily evaluated) image, it can be used as input for another on-demand operation. We can pass the above intensity corrected image as the input argument

to a cell loading function that uses a Gaussian convolution Op. The result is again a lazy image. When a pixel of this image is accessed, it will trigger the Gaussian cell loader to produce the cell containing the requested pixel. This will access pixels of its input, which will in turn trigger the intensity correction cell loader to produce the cell(s) containing those pixels.

4.2 Performance

A common pattern in software is to have a slow-but-general case, and a fast special case. Consider writing a method to create a new copy of an image. If we were to type its parameters on a general image data interface, we obtain the flexibility and extensibility to copy any sort of image; the actual data could be stored by column or by row, on GPU or in memory. If we were instead to restrict our method to images backed by primitive Java byte arrays, we could utilize that information to copy much faster (using `java.lang.System.arraycopy`). To provide the benefits of both choices, many programs employ case logic: if the image data is backed by an array in memory, copy it the “fast way”, otherwise copy it the “slow but general” way element by element. Unfortunately, this logic not only makes programs more complex and difficult to maintain, it also prevents adding new cases without changing the source code. If we later want to add a case that copies image data between two buffers on a graphical processing unit (GPU) but we don’t have access to the original source code, we must write a new method. In such situations, plugin-based frameworks like SciJava Ops allow developers to simply add another plugin to operate upon GPU buffers, and the matcher will take care of selecting this Op in cases where the image data type indicates it



would be compatible. This “extensible case logic” allows SciJava Ops to take advantage of future performance improvements with minimal maintenance.

While the SciJava Ops matcher allows future performance improvements, it is vital that the matcher must not impact performance itself. To quantify framework overhead, we defined a simple Op `benchmark.increment(data: byte[])`, defined as a static Java method, to increment the first byte in the provided array. This behavior was chosen to minimize the method’s runtime so that runtime consists almost entirely of overhead. This Op was called both through direct invocation, and through the SciJava Ops framework. Each execution pathway was benchmarked using the Java Microbenchmark Harness (JMH) (<https://github.com/openjdk/jmh>) on an AMD Ryzen 7 3700X CPU and 32 GB of 3,200 MHz DDR4 memory, and summarized with average execution times (second per invocation). Figure 11 visualizes the mean execution overhead of each profiled method, with error bars denoting the observed range, using both a logarithmic (left) and linear (right) scale. Direct (static method) invocation results show that our Op indeed imposes a negligible runtime of around 1 microsecond, ensuring that any

time spent in other benchmarks are solely attributable to the SciJava Ops and ImageJ Ops frameworks. Basic execution via SciJava Ops imposes an overhead of approximately 100 microseconds, however by using an internal cache SciJava Ops can reduce the overhead of repeated requests to approximately 3 microseconds. These results suggest that most of SciJava Ops’s overhead comes from the matching process. As we would expect, the framework features of Op adaptation and parameter conversion provide additional overhead, and their composition shows that this overhead is additive. Perhaps most importantly, we show that the overhead imposed by SciJava Ops is significantly less than that of ImageJ Ops, with direct matching running approximately 70 times faster than its predecessor.

4.3 Usability

To solve real-world problems, users often need to combine tools from multiple libraries into more complex workflows; SciJava Ops’s framework features make such synergy easier to achieve on several fronts:

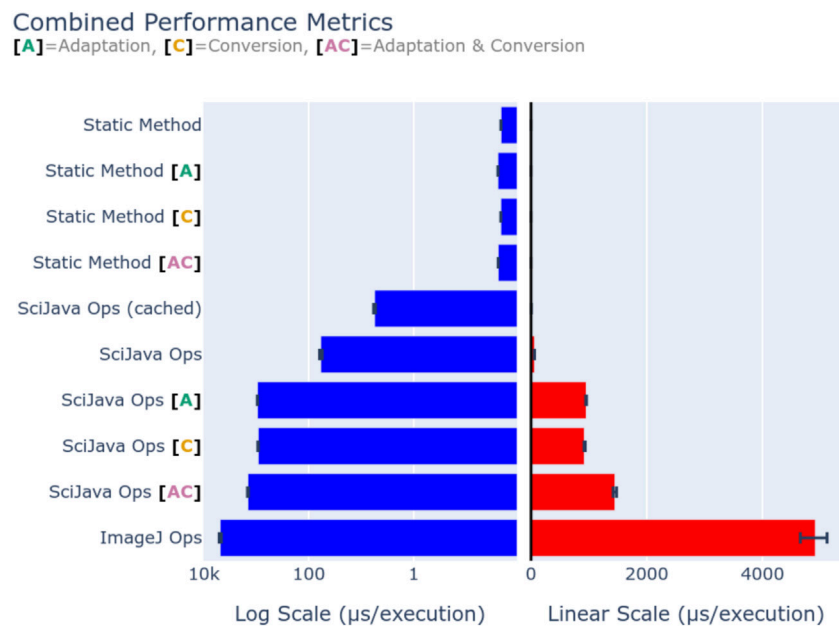


FIGURE 11

Execution overhead (in microseconds) of the Op benchmark `increment (data: byte[])`, through `static method` invocation, `SciJava Ops`, and `ImageJ Ops`. Op adaptation [A] was additionally benchmarked for `SciJava Ops` by requesting the Op as a function Op, and for static invocation by creating an output analogous to the Op adaptation routine. Parameter conversion [C] was benchmarked for `SciJava Ops` by requesting the Op with a `double []` parameter, and conversion was performed manually for static invocation for comparison. Finally, a composition of adaptation and conversion [A + C] was benchmarked to profile the overhead of their combination. In all `SciJava Ops` benchmarks the Op cache was disabled to ensure the performance of the Op matcher was measured, except in the case where caching was explicitly enabled for the sake of comparison to static invocation. Execution overhead is displayed on both a logarithmic (left) and linear (right) scale for comparison.

4.3.1 Unified documentation and access

Multiple libraries each have their own locations for documentation, which then must be cross-referenced to compare similar algorithms. When using `SciJava Ops` the documentation for each algorithm present can be accessed consistently through the `OpEnvironment.help` API, optionally including technical comments and type descriptions. Furthermore, the declarative nature of `SciJava Ops` means when multiple libraries contain similar algorithms, they can be classified and discovered using the same name. Finally, standardization of algorithms as function, computer and inplace Ops ensures accessing algorithms across varied libraries can always follow the same pattern - name, inputs, outputs - even when, for example, `OpenCV` algorithms declare their preallocated output buffers as their first parameter.

4.3.2 Automatic, extensible type conversion

Without `SciJava Ops`, combining algorithm libraries requires the addition of “glue” layers such as `PyImageJ` and `ImageJ-OpenCV` to convert between analogous data structures, and the user is responsible for managing that conversion manually. `SciJava Ops` removes the need for this manual effort, allowing unified library access that is seamless from the user’s perspective: they don’t need to consider where an algorithm implementation comes from to benefit from it. As demonstrated in Section 4.1, at the core of this work are the `engine.convert` Ops which are visible only to developers, but invoked transparently on demand. By avoiding these manual steps, users can save time constructing their workflows and

accomplish real world tasks faster, with fewer opportunities for error.

4.3.3 Transparent algorithm improvements over time

We also designed `SciJava Ops` to adapt to users’ needs over time: as new algorithms are developed promising significant improvements over the previous generation, users are faced with learning an entirely new library, with its own minutiae, documentation sources, and possibly even new data structures or programming languages. `SciJava Ops` eases these transitions through its declarative nature, as improved algorithms can simply be added as new Ops as they are developed. If they are defined with the same name and accept the same parameters as an existing Op, future requests using that name and parameters will instead match the new and improved Op, allowing users to enjoy the benefits of improved performance or behavior without changing their existing workflows. Even in cases where changes are required, the consistency of form and access in `SciJava Ops` empowers users to work in a familiar environment, with reliable API.

4.4 Limitations

4.4.1 Op reproducibility is not enforceable

With the `SciJava Ops` `Image` library, we have made a concerted effort to design each Op to behave deterministically and reproducibly, yielding the same result when given the same

arguments (parameter values). But some algorithms are non-deterministic, with varying results upon multiple executions with the same inputs. This can happen when an algorithm uses a random number generator without a fixed seed value, or stores persistent data that affects the computation across executions, or uses multithreaded computation in a way that accumulates results differently depending on the timing of each thread's calculations. Ideally, such algorithms would be adjusted to avoid such non-deterministic behavior: use a fixed or user-provided seed for random number generation, accept all data as explicit inputs to the Op, and carefully coordinate all threads of a multithreaded computation to ensure reproducible calculation. But Ops has no mechanism to ensure that Ops-based analyses always produce the same results across computing environments, which would better foster reproducible science.

4.4.2 Converting arguments between types may require copying data

Parameter conversion incurs significant overhead whenever argument data must be copied. To that end, built-in `engine.convert` Ops wrap data whenever feasible, yet many conversions are infeasible without explicit data copying. For example, while `ImgLib2` image structures allow zero-copy conversion of `OpenCV Mat` objects, the `Mat` interface provides no opportunity for zero-copy conversion in the other direction. Parameter conversion still allows SciJava Ops to provide seamless interoperability, but forced copies impact overall performance.

4.4.3 The Op engine discriminates only on parameter types, not values

With the current architecture of SciJava Ops, it is not possible to have two Ops with the same name, each operating on a different subset of the same input types. Consider the case of two-dimensional convolution; depending on kernel size, arithmetic convolution may or may not be faster than addition within the frequency domain. Since a kernel's size cannot be inferred from its type, the SciJava Ops framework cannot delegate between these two algorithms. Instead, delegation would require a single Op to wrap both algorithms with conditional logic, limiting extensibility.

4.4.4 Op executions are complicated to trace

While the SciJava Ops Framework is built to foster reproducible workflows, it can be difficult to determine which blocks of code correspond to each workflow component. Some ambiguity comes from the declarative nature of Op requests, which necessarily adds a layer of indirection. Furthermore, Op adaptation, parameter conversion and other internal engine features further obfuscate the underlying Op through additional transformations. To obtain an understanding of which Ops were run, two different components of the SciJava Ops framework can be used. Each Op contains an `InfoTree` data structure, which provides a full graph of dependencies, adaptations and parameter conversions. Additionally, the `OpHistory` keeps a record of each Op that edited an output buffer, enabling reconstruction of the set of Ops used to produce a given Object. While we do not predict SciJava Ops users to need access to these features often, it is nonetheless an extra level of complication compared to direct imperative function calls.

4.4.5 The Ops engine runs on the Java Virtual Machine

The JVM is a powerful piece of technology, offering good performance and a well-designed type system. But the OpenJDK platform providing the JVM can weigh in at hundreds of megabytes in size, and may be undesirable as a dependency in scenarios with limited space such as embedded systems. While the architecture of SciJava Ops is not fundamentally tied to Java or the JVM specifically, the current implementation is, and more effort would be necessary to change that—e.g. implementing another Ops engine on top of a different platform, or adjusting the Ops codebase to work with an Ahead-of-Time compiler such as Kotlin Native (<https://kotlinlang.org/docs/native-overview.html>) or GraalVM's Native Image (<https://www.graalvm.org/jdk21/reference-manual/native-image/>).

4.4.6 Naming is hard

As the library of available Ops grows, care must be taken to ensure consistent naming across different collections of Ops. While there is no requirement that two Ops with the same name produce binary-identical results when passed the same inputs—and indeed such a requirement would be very challenging to achieve in practice—it is a judgment call whether two pieces of code performing variations of the same idea are close enough to be considered the “same algorithm” and therefore be labeled with the same Op name. Like any ontology project, guidelines will need to emerge and collections will need to be well curated in order to prevent the growth of an inconsistent mess over time.

4.5 Future directions

The current state of SciJava Ops provides the core frameworks for extensibility, interoperability and reproducibility, with a number of standard image processing algorithms from SciJava Ops Image, `ImgLib2`, and `OpenCV` bindings. This is all in service of progress towards a single unified cross-platform algorithm entry point. We believe the next steps in this path are as follows.

4.5.1 Wrapping new libraries from python, C++, and other software languages

There are many existing libraries that, when made useable from SciJava Ops, will further expand workflow possibilities. Many gaps in SciJava Ops Image can be filled using libraries built upon the Python scientific stack, such as `SciPy`, `scikit-image`, `scikit-learn`, and `PyTorch`. C++ libraries built atop `CUDA` and GPUs can be wrapped to increase performance over existing algorithms. The result will be full access to the most popular algorithms from these communities, all using the same syntactically consistent Op builder pattern, facilitating more seamless combination of algorithms from a diversity of sources.

4.5.2 Integrate SciJava Ops into other scientific applications

In this text we have highlighted Fiji as a natural entry point for SciJava Ops, but as we have shown SciJava Ops can be used in environments outside of Fiji as well. Beyond the current Python support powered by `scjjava`, users can also write standalone scripts from any Java-based scripting language (e.g., `Groovy`, `Jython` or `Kotlin`).

We plan to continue working with the imaging software community to integrate SciJava Ops into additional image analysis tools (e.g., napari, CellProfiler, Icy) so that the same algorithms can be used the same way across applications, which will not only bring value to the users of these tools, but also encourage the future integration of additional algorithms libraries into Ops as an easier route to dissemination.

4.5.3 An algorithms index for algorithm discovery and exposure

Today there is already a question of “what image processing algorithms are available?” and simply expanding SciJava Ops to integrate more algorithm libraries will not answer this question by itself. We envision a unified web-based algorithm repository, with clear routes to identifying which algorithms are available and which library (ies) implement said algorithm, to maximize accessibility. Developers can browse such an index to determine whether their algorithm has already been implemented in another project and to calculate the comparative burden of wrapping an existing algorithm instead of reimplementing it within their own library. They can also publish to such an index to expose their new implementations, avoiding reimplementation done purely because another developer could not find any existing code. Users can browse such an index to determine which library might best suit their needs, rather than extensively searching through the documentation of many different packages.

We believe that SciJava Ops can ultimately be leveraged to build such an algorithm repository, providing users and developers with both the means to locate algorithms of interest and also practical instructions for including them in the workflow environments they prefer. The index of algorithm implementations will be automatically generated and maintained from an aggregated collection of known Ops, while also featuring a general page about each algorithm (e.g., filter.gauss) with a community-curated wiki-style introductions to the algorithm along with links to all known implementations across various software packages.

5 Conclusion

SciJava Ops represents the culmination of our experience creating extensible frameworks for use in scientific communities. Designed as a foundational layer for algorithms across the scientific image analysis community, it aims to meet technical requirements across many segments of a broad target audience: strong typing guarantees for developers, declarative calls for users, extensible code-free integration paths coupled with easy-to-use annotations, with automatic adaptation and type conversion to aid longevity and maintainability of any algorithm library in our ecosystem. We believe we have succeeded at creating a framework capable of bridging the gaps between libraries of algorithms in Java and beyond, facilitating accessible computation with less time lost to recapitulation and fragmentation. We are excited for the potential for growth and contributions as SciJava Ops enters into public use and continues to grow, approaching our vision of a consolidated image processing ecosystem, where users can discover and harness algorithms across many software technologies in a unified way.

6 Materials and methods

6.1 RLTV deconvolution and python (ScyJava) datasets

HeLa and 3T3 cells were plated in 8-well no. 1.5H glass bottom slides (Ibidi) and maintained at 37°C, ~50% humidity, and 5% CO₂ prior to fixation in 4% paraformaldehyde in PBS. Both HeLa and 3T3 fixed cells were stained with 4',6-diamidino-2-phenylindole (DAPI). Z-stack images with a step size of 0.1 μm were acquired on a Nikon Ti-Eclipse inverted wide-field epifluorescent microscope (Nikon Corporation) using a 1.45NA 100x Plan Apo oil immersion objective lens and an ORCA-Flash4.0 CMOS camera (Hamamatsu Photonics, Skokie, IL USA). DAPI images were acquired with a 325–375 nm emission/435–485 nm excitation filter set.

6.2 SACA dataset

HeLa cells were in 24-well glass bottom plates (Cellvis) and fixed in 4% paraformaldehyde and stained with DAPI 48 h post transfection with a dual color fluorescent HIV_{NL4-3} construct expressing Gag-YFP and MS2-mCherry fusion proteins. Images were acquired on an A1R laser scanning confocal microscope built on Nikon Ti-Eclipse inverted base using a 1.4NA 60x Plan Apo oil immersion objective lens. DAPI, YFP and mCherry were excited by the 405 nm, 488 nm and 561 nm diode lasers respectively.

6.3 FLIM dataset and analysis

Fixed stained slides of bovine pulmonary artery endothelial cells (BPAEC) were obtained from ThermoFisher Scientific (FluoCells™ Prepared Slide #1, CAT#F36924). The sample contains MitoTracker™ Red CMXRos, Alexa Fluor™ 488 Phalloidin, and DAPI nuclear stains and further details can be obtained from the manufacturer. The sample was imaged using a 40xWI 1.15NA Nikon objective lens with multiphoton excitation at 740 nm and collected with a 650 nm shortpass filter to image all three fluorophores. The imaging and scanning was controlled by Openscan-LSM (<https://github.com/openscan-lsm/OpenScan>) and SPC180 (Becker-Hickl GmbH, Berlin, Germany) fast timing electronics card. The FLIM data was modeled using a mono-exponential decay and Levenberg-Marquardt algorithm (LMA) fitting using SciJava Ops FLIM.

Data availability statement

The datasets generated and analyzed for this study can be found on the SciJava Ops website (<https://ops.scijava.org/en/1.1.0/paper-2024>).

Author contributions

GS: Conceptualization, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing—original draft, Writing—review and editing. CR: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing—original draft, Writing—review and editing. MH: Formal Analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing—original draft, Writing—review and editing. EE: Data curation, Resources, Software, Validation, Visualization, Writing—original draft, Writing—review and editing. DK: Conceptualization, Investigation, Software, Writing—review and editing. MW: Conceptualization, Investigation, Software, Writing—review and editing. CB: Conceptualization, Investigation, Software, Writing—review and editing. T-OB: Software, Writing—review and editing. SH: Software, Writing—review and editing. BN: Software, Writing—review and editing. AW: Software, Writing—review and editing. JS: Conceptualization, Software, Writing—review and editing. TP: Writing—review and editing. SS: Writing—review and editing. MB: Supervision, Writing—review and editing, Funding acquisition, Project administration. KE: Funding acquisition, Project administration, Resources, Supervision, Writing—review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. We acknowledge funding support from the Chan Zuckerberg Initiative (CR and KE), NIH Grants P41GM135019 (KE) and U54CA268069 (KE) as well as additional support from the Morgridge Institute for Research (EE and KE).

References

- Ahlers, J., Althviz Moré, D., Amsalem, O., Anderson, A., Bokota, G., Boone, P., et al. (2023). napari: a multi-dimensional image viewer for Python. Available at: <https://zenodo.org/record/3555620> (Accessed April 4, 2024).
- Arena, E. T., Rueden, C. T., Hiner, M. C., Wang, S., Yuan, M., and Eliceiri, K. W. (2017). Quantitating the cell: turning images into numbers with ImageJ. *WIREs Dev. Biol.* 6 (2), e260. doi:10.1002/wdev.260
- Arshadi, C., Günther, U., Eddison, M., Harrington, K. I. S., and Ferreira, T. A. (2021). SNT: a unifying toolbox for quantification of neuronal anatomy. *Nat. Methods* 18 (4), 374–377. doi:10.1038/s41592-021-01105-7
- Arzt, M., Deschamps, J., Schmied, C., Pietzsch, T., Schmidt, D., Tomancak, P., et al. (2022). LABKIT: labeling and segmentation toolkit for big image data. *Front. Comput. Sci.* 4, 777728. doi:10.3389/fcomp.2022.777728
- Bankhead, P., Loughrey, M. B., Fernández, J. A., Dombrowski, Y., McArt, D. G., Dunne, P. D., et al. (2017). QuPath: open source software for digital pathology image analysis. *Sci. Rep.* 7 (1), 16878. doi:10.1038/s41598-017-17204-5
- Bertero, M., and Boccacci, P. (2005). A simple method for the reduction of boundary effects in the Richardson-Lucy approach to image deconvolution. *Astron Astrophys.* 437 (1), 369–374. doi:10.1051/0004-6361:20052717
- Berthold, M. R., Cebron, N., Dill, F., Gabriel, T. R., Kötter, T., Meinel, T., et al. (2008). “KNIME: the konstanz information miner,” in *Data analysis, machine learning and applications*. Editors C. Preisach, P. D. H. Burkhardt, P. D. L. Schmidt-Thieme, and P. D. R. Decker (Germany: Springer Berlin Heidelberg), 319–326. Available at: http://link.springer.com/chapter/10.1007/978-3-540-78246-9_38.
- Biggs, D. S. C., and Andrews, M. (1997). Acceleration of iterative image restoration algorithms. *Appl. Opt.* 36 (8), 1766. doi:10.1364/ao.36.001766
- Bradski, G. (2000). The OpenCV library. *Dr. Dobbs J.* 25 (11), 120–125. doi:10.3233/978-1-60750-929-5-586
- Cardona, A., Saalfeld, S., Schindelin, J., Arganda-Carreras, I., Preibisch, S., Longair, M., et al. (2012). TrakEM2 software for neural circuit reconstruction. *PLoS ONE* 7 (6), e38011. doi:10.1371/journal.pone.0038011
- De Chaumont, F., Dallongeville, S., Chenouard, N., Hervé, N., Pop, S., Provoost, T., et al. (2012). Icy: an open bioimage informatics platform for extended reproducible research. *Nat. Methods* 9 (7), 690–696. doi:10.1038/nmeth.2075
- Dey, N., Blanc-Feraud, L., Zimmer, C., Roux, P., Kam, Z., Olivo-Marin, J. C., et al. (2006). Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution. *Microsc. Res. Tech.* 69 (4), 260–266. doi:10.1002/jemt.20294
- Dietz, C., and Berthold, M. R. (2016). “KNIME for open-source bioimage analysis: a tutorial,” in *Focus on bio-image informatics*. Editors W. H. De Vos, S. Munck, and J. P. Timmermans (Cham: Springer International Publishing). Available at: http://link.springer.com/10.1007/978-3-319-28549-8_7.
- Dietz, C., Rueden, C. T., Helfrich, S., Dobson, E. T. A., Horn, M., Eglinger, J., et al. (2020). Integration of the ImageJ ecosystem in the KNIME Analytics platform. *Front. Comput. Sci.* 2, 8. doi:10.3389/fcomp.2020.00008
- Dobson, E. T. A., Cimini, B., Klemm, A. H., Wählby, C., Carpenter, A. E., and Eliceiri, K. W. (2021). ImageJ and CellProfiler: complements in open-source bioimage analysis. *Curr. Protoc.* 1 (5), e89. doi:10.1002/cpz1.89
- Domander, R., Felder, A. A., and Doube, M. (2021). BoneJ2 - refactoring established research software. *Wellcome Open Res.* 6, 37. doi:10.12688/wellcomeopenres.16619.2
- Dominguez, C., Heras, J., and Pascual, V. (2017). IJ-OpenCV: combining ImageJ and OpenCV for processing images in biomedicine. *Comput. Biol. Med.* 84, 189–194. doi:10.1016/j.compbiomed.2017.03.027

Acknowledgments

The authors acknowledge Barry DeZonia for his initial effort to develop a functional image processing library for ImgLib2; Martin Horn for co-founding the ImageJ Ops project; Stephan Preibisch for co-developing the ImgLib2 library; Deborah Schmidt for helpful ideas and discussions; Jenu Chacko for FLIM data and analysis guidance; Dasong Gao for his work on the FLIMJ Ops; Sydney Lesko for confocal data for SACA; and all the software developers who contributed to Ops over the years.

Conflict of interest

Authors DK, MW, CB, SH, AW, and MB were employed by KNIME GmbH. Author JS was employed by Microsoft Corporation. Author BN was employed by True North Intelligent Algorithms.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Ershov, D., Phan, M. S., Pylvänäinen, J. W., Rigaud, S. U., Le Blanc, L., Charles-Orszag, A., et al. (2022). TrackMate 7: integrating state-of-the-art segmentation algorithms into tracking pipelines. *Nat. Methods* 19 (7), 829–832. doi:10.1038/s41592-022-01507-1
- Gao, D., Barber, P. R., Chacko, J. V., Sagar, M. A. K., Rueden, C. T., Grisli, A. R., et al. (2020). FLIMJ: an open-source ImageJ toolkit for fluorescence lifetime image data analysis. *PLoS ONE* 15 (12), e0238327. doi:10.1371/journal.pone.0238327
- Gibson, S. F., and Lanni, F. (1991). Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy. *JOSA A* 8 (10), 1601–1613. doi:10.1364/josaa.8.001601
- Haase, R., Royer, L. A., Steinbach, P., Schmidt, D., Dibrov, A., Schmidt, U., et al. (2020). CLIJ: GPU-accelerated image processing for everyone. *Nat. Methods* 17 (1), 5–6. doi:10.1038/s41592-019-0650-1
- Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., et al. (2020). Array programming with NumPy. *Nature* 585 (7825), 357–362. doi:10.1038/s41586-020-2649-2
- Hiner, M. C., Rueden, C. T., and Eliceiri, K. W. (2016). SCIFIO: an extensible framework to support scientific image formats. *BMC Bioinforma.* 17 (1), 521. doi:10.1186/s12859-016-1383-0
- Hiner, M. C., Rueden, C. T., and Eliceiri, K. W. (2017). ImageJ-MATLAB: a bidirectional framework for scientific image analysis interoperability. *Bioinformatics* 33 (4), 629–630. doi:10.1093/bioinformatics/btw681
- Hörl, D., Rojas Rusak, F., Preusser, F., Tillberg, P., Randel, N., Chhetri, R. K., et al. (2019). BigStitcher: reconstructing high-resolution image datasets of cleared and expanded samples. *Nat. Methods* 16 (9), 870–874. doi:10.1038/s41592-019-0501-0
- Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., et al. (2016). in *Jupyter Notebooks – a publishing format for reproducible computational workflows*. Editors F. Loizides and B. Schmidt (Germany: IOS Press), 87–90. Available at: <https://eprints.soton.ac.uk/403913/>.
- Lambert, T., and Waters, J. (2023). Towards effective adoption of novel image analysis methods. *Nat. Methods* 20 (7), 971–972. doi:10.1038/s41592-023-01910-2
- Legland, D., Arganda-Carreras, I., and Andrey, P. (2016). MorphoLibJ: integrated library and plugins for mathematical morphology with ImageJ. *Bioinformatics* 32 (22), 3532–3534. doi:10.1093/bioinformatics/btw413
- Li, Q., Lau, A., Morris, T. J., Guo, L., Fordyce, C. B., and Stanley, E. F. (2004). A syntaxin 1, α_9 , and N-type calcium channel complex at a presynaptic nerve terminal: analysis by quantitative immunocolocalization. *J. Neurosci. Off. J. Soc. Neurosci.* 24 (16), 4070–4081. doi:10.1523/jneurosci.0346-04.2004
- Li, X., Zhang, Y., Wu, J., and Dai, Q. (2023). Challenges and opportunities in bioimage analysis. *Nat. Methods* 20 (7), 958–961. doi:10.1038/s41592-023-01900-4
- Linkert, M., Rueden, C. T., Allan, C., Burel, J. M., Moore, W., Patterson, A., et al. (2010). Metadata matters: access to image data in the real world. *J. Cell. Biol.* 189 (5), 777–782. doi:10.1083/jcb.201004104
- Lowe, B. C., Chen, D. T., Ibáñez, L., and Blezek, D. (2013). The design of SimpleITK. *Front. Neuroinformatics* 7, 45. doi:10.3389/fninf.2013.00045
- Manders, E. M. M., Verbeek, F. J., and Aten, J. A. (1993). Measurement of colocalization of objects in dual-colour confocal images. *J. Microsc.* 169 (3), 375–382. doi:10.1111/j.1365-2818.1993.tb03313.x
- Moore, J., Allan, C., Besson, S., Burel, J. M., Diel, E., Gault, D., et al. (2021). OME-NGFF: a next-generation file format for expanding bioimaging data-access strategies. *Nat. Methods* 18 (12), 1496–1498. doi:10.1038/s41592-021-01326-w
- Moore, J., Linkert, M., Blackburn, C., Carroll, M., Ferguson, R. K., Flynn, H., et al. (2015). “OMERO and Bio-Formats 5: flexible access to large bioimaging datasets at scale,” in *Medical imaging 2015: image processing* (USA: SPIE), 37–42. Available at: <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/9413/941307/OMERO-and-Bio-Formats-5-flexible-access-to-large/10.1117/12.2086370.full>.
- Nogare, D. D., Hartley, M., Deschamps, J., Ellenberg, J., and Jug, F. (2023). Using AI in bioimage analysis to elevate the rate of scientific discovery as a community. *Nat. Methods* 20 (7), 973–975. doi:10.1038/s41592-023-01929-5
- Ouyang, W., Beuttenmueller, F., Gómez-de-Mariscal, E., Pape, C., Burke, T., García-López-de-Haro, C., et al. (2022). BioImage model Zoo: a community-driven resource for accessible deep learning in BioImage analysis. Available at: <http://biorxiv.org/lookup/doi/10.1101/2022.06.07.495102> (Accessed April 4, 2024).
- Ouyang, W., Eliceiri, K. W., and Cimini, B. A. (2023). Moving beyond the desktop: prospects for practical bioimage analysis via the web. *Front. Bioinforma.* 3, 1233748. doi:10.3389/fbinf.2023.1233748
- Ouyang, W., Mueller, F., Hjeltnar, M., Lundberg, E., and Zimmer, C. (2019). ImJoy: an open-source computational platform for the deep learning era. *Nat. Methods* 16 (12), 1199–1200. doi:10.1038/s41592-019-0627-0
- Pietzsch, T., Preibisch, S., Tomančák, P., and Saalfeld, S. (2012). ImgLib2—generic image processing in Java. *Bioinformatics* 28 (22), 3009–3011. doi:10.1093/bioinformatics/bts543
- Rubens, U., Hoyoux, R., Vanosmael, L., Ouras, M., Tasset, M., Hamilton, C., et al. (2019). Cytomine: toward an open and collaborative software platform for digital pathology bridged to molecular investigations. *PROTEOMICS – Clin. Appl.* 13 (1), 1800057. doi:10.1002/prca.201800057
- Rubens, U., Mormont, R., Paavolainen, L., Bäcker, V., Pavie, B., Scholz, L. A., et al. (2020). BIAFLOWS: a collaborative framework to reproducibly deploy and benchmark bioimage analysis workflows. *Patterns* 1 (3), 100040. doi:10.1016/j.patter.2020.100040
- Rueden, C. T., Hiner, M. C., Evans, E. L., Pinkert, M. A., Lucas, A. M., Carpenter, A. E., et al. (2022). PyImageJ: a library for integrating ImageJ and Python. *Nat. Methods* 19 (11), 1326–1327. doi:10.1038/s41592-022-01655-4
- Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., et al. (2017). ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinforma.* 18 (1), 529. doi:10.1186/s12859-017-1934-z
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9 (7), 676–682. doi:10.1038/nmeth.2019
- Schindelin, J., Rueden, C. T., Hiner, M. C., and Eliceiri, K. W. (2015). The ImageJ ecosystem: an open platform for biomedical image analysis. *Mol. Reprod. Dev.* 82 (7–8), 518–529. doi:10.1002/mrd.22489
- Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9 (7), 671–675. doi:10.1038/nmeth.2089
- Schroeder, W. J., Avila, L. S., and Hoffman, W. (2000). Visualizing with VTK: a tutorial. *IEEE Comput. Graph Appl.* 20 (5), 20–27. doi:10.1109/38.865875
- Selzer, G. J., Rueden, C. T., Hiner, M. C., Evans, E. L., Harrington, K. I. S., and Eliceiri, K. W. (2023). napari-imagej: ImageJ ecosystem access from napari. *Nat. Methods* 20 (10), 1443–1444. doi:10.1038/s41592-023-01990-0
- Stirling, D. R., Swain-Bowden, M. J., Lucas, A. M., Carpenter, A. E., Cimini, B. A., and Goodman, A. (2021). CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinforma.* 22 (1), 433. doi:10.1186/s12859-021-04344-9
- TCSPC (2021). *The bh TCSPC Handbook*. 9th edition. China: Becker and Hickl GmbH. Available at: <https://www.becker-hickl.com/literature/documents/flim/the-bh-tcspc-handbook-9th-edition-2021/>.
- van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., et al. (2014). scikit-image: image processing in Python. *PeerJ* 2, e453. doi:10.7717/peerj.453
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* 17 (3), 261–272. doi:10.1038/s41592-019-0686-2
- Wang, S., Arena, E. T., Becker, J. T., Bement, W. M., Sherer, N. M., Eliceiri, K. W., et al. (2019). Spatially adaptive colocalization analysis in dual-color fluorescence microscopy. *IEEE Trans. Image Process. Publ. IEEE Signal Process. Soc.* 28, 4471–4485. doi:10.1109/tip.2019.2909194
- Wiki, J. (2024). Deconvolution. Available at: <https://imagej.github.io/imaging/deconvolution>.
- Yaniv, Z., Lowekamp, B. C., Johnson, H. J., and Beare, R. (2018). SimpleITK image-analysis notebooks: a collaborative environment for education and reproducible research. *J. Digit. Imaging* 31 (3), 290–303. doi:10.1007/s10278-017-0037-8
- Yoo, T. S., Ackerman, M. J., Lorensen, W. E., Schroeder, W., Chalana, V., Aylward, S., et al. (2002). Engineering and algorithm design for an image processing API: a technical report on ITK – the insight toolkit. *Med. Meets Virtual Real.* 0210. . doi:10.3233/978-1-60750-929-5-586