

# AI in Microscopy: A Biolmaging Guide

# Table of contents

<b>Welcome</b>	<b>8</b>
<b>1 Preface</b>	<b>9</b>
1.1 Introduction . . . . .	9
1.2 Textbook Outline . . . . .	9
<b>I Getting Started with AI</b>	<b>12</b>
<b>2 AI Primer</b>	<b>13</b>
2.1 Include section headers as appropriate . . . . .	13
2.1.1 Sub-subsection headers are also available . . . . .	13
2.2 Bibliography and Citations . . . . .	14
2.3 Adding to the Glossary . . . . .	14
2.4 Code and Equations . . . . .	14
2.4.1 Embedding Figures . . . . .	16
2.5 Quarto has additional features. . . . .	18
<b>3 Large Language Models (LLMs)</b>	<b>19</b>
3.1 Foundations of Large Language Models . . . . .	19
3.2 Multi-modal AI: Vision-Language Models and Generative AI . . . . .	19
3.3 AI Agents for Microscopy Workflows . . . . .	20
3.4 Challenges and Limitations . . . . .	20
3.5 Future Directions . . . . .	20
3.6 Practical Guide: Getting Started with LLMs for Microscopy . . . . .	20
<b>4 Architectures and Loss Models</b>	<b>21</b>
4.1 Include section headers as appropriate . . . . .	21
4.1.1 Sub-subsection headers are also available . . . . .	21
4.2 Bibliography and Citations . . . . .	21
4.3 Adding to the Glossary . . . . .	22
4.4 Code and Equations . . . . .	22
4.4.1 Embedding Figures . . . . .	24
4.5 Quarto has additional features. . . . .	25

<b>II Image Acquisition</b>	<b>26</b>
<b>5 Collecting Training Data</b>	<b>27</b>
5.1 Include section headers as appropriate . . . . .	27
5.1.1 Sub-subsection headers are also available . . . . .	27
5.2 Bibliography and Citations . . . . .	27
5.3 Adding to the Glossary . . . . .	28
5.4 Code and Equations . . . . .	28
5.4.1 Embedding Figures . . . . .	30
5.5 Quarto has additional features. . . . .	31
<b>6 Image Restoration</b>	<b>32</b>
6.1 General Concepts in Image Restoration . . . . .	33
6.1.1 Image Degradation Model . . . . .	33
6.1.2 Image Restoration with Traditional Approaches . . . . .	34
6.1.3 Image Restoration with Deep Learning-Based Approaches . . . . .	37
6.2 Deep Learning-Based Techniques for Image Restoration . . . . .	38
6.2.1 Denoising . . . . .	38
6.2.2 Deconvolution . . . . .	39
6.2.3 Deaberration . . . . .	42
6.2.4 Resolution Enhancement . . . . .	42
6.3 Practical Guidelines for Image Restoration . . . . .	44
6.3.1 Considerations for Dataset Preparation . . . . .	44
6.3.2 Choosing a Network Architecture . . . . .	46
6.3.3 Training . . . . .	46
6.3.4 Validation and Deployment . . . . .	47
6.4 Limitations and Future Perspectives . . . . .	47
6.4.1 Caveats of AI-based image restoration . . . . .	47
6.4.2 Outlook for the Future . . . . .	48
<b>7 Adding AI to Your Hardware</b>	<b>50</b>
<b>III Image Analysis</b>	<b>52</b>
<b>8 How do you select and find a tool?</b>	<b>53</b>
8.1 Introduction . . . . .	53
8.2 Assessing your requirements for a tool (in a perfect world) . . . . .	53
8.2.1 Questions about your overall workflow: . . . . .	55
8.2.2 Questions about your resources: time, expertise, and compute: . . . . .	55
8.2.3 Questions about your input data: . . . . .	56
8.3 Assessing your requirements for a tool (in the real world) . . . . .	58
8.4 How do I decide when my workflow is “good enough”? . . . . .	59

8.5	Choosing the kind of tool to use . . . . .	61
8.6	Finding tools and models . . . . .	66
8.6.1	Where to look for tools and models . . . . .	66
8.7	Case Studies . . . . .	67
8.7.1	Case Study 1: . . . . .	67
8.7.2	Case Study 2: . . . . .	71
8.8	Conclusion . . . . .	76
<b>9</b>	<b>How to Train and Use Deep Learning Models in Microscopy</b>	<b>77</b>
9.1	Introduction . . . . .	77
9.1.1	The challenge . . . . .	77
9.2	Preparing for your project . . . . .	78
9.2.1	Defining your task and success criteria . . . . .	78
9.2.2	Evaluating alternatives: Is DL the right choice? . . . . .	79
9.3	Implementing a DL segmentation workflow . . . . .	81
9.3.1	Overview of a typical DL segmentation workflow . . . . .	81
9.3.2	Selecting a suitable DL approach . . . . .	82
9.3.3	Deciding whether to train a new model . . . . .	82
9.3.4	Preparing your dataset for training . . . . .	84
9.3.5	Training a segmentation model from scratch . . . . .	87
9.3.6	Fine-Tuning Pre-existing Models . . . . .	90
9.3.7	Evaluating the performance of your model . . . . .	92
9.3.8	Deploying your model on new data . . . . .	95
9.3.9	Troubleshooting Common Problems . . . . .	95
9.4	Further considerations for DL segmentation . . . . .	99
9.4.1	Choosing the Right Tools for DL . . . . .	99
9.4.2	Managing Computational Resources . . . . .	100
9.4.3	Ensuring Reproducibility in DL . . . . .	101
9.5	Summary & Outlook . . . . .	102
<b>10</b>	<b>Output Quality</b>	<b>104</b>
10.1	Metrics and Losses . . . . .	104
10.2	What Metric to Pick? . . . . .	105
10.3	Examples . . . . .	105
10.3.1	Segmentation Metrics . . . . .	105
10.3.2	Cell Tracking Metrics . . . . .	105
<b>11</b>	<b>Outlook</b>	<b>106</b>
11.1	Include section headers as appropriate . . . . .	106
11.1.1	Sub-subsection headers are also available . . . . .	106
11.2	Bibliography and Citations . . . . .	106
11.3	Adding to the Glossary . . . . .	107

11.4 Code and Equations . . . . .	107
11.4.1 Embedding Figures . . . . .	109
11.5 Quarto has additional features. . . . .	110

<b>Glossary</b>	<b>111</b>
Aberration . . . . .	111
Adaptive Optics . . . . .	111
Autoencoder . . . . .	111
Backpropagation . . . . .	111
Batch . . . . .	111
Bayesian Optimization . . . . .	112
Binary Segmentation . . . . .	112
CARE (Content-aware image restoration) . . . . .	112
Computer Vision . . . . .	112
Convolution . . . . .	112
Convolutional Neural Networks (CNNs) . . . . .	112
Data Augmentation . . . . .	113
Deconvolution . . . . .	113
Domain Randomization . . . . .	113
Effect Size . . . . .	113
Epoch . . . . .	113
F1 Score . . . . .	113
False Negatives . . . . .	114
False Positives . . . . .	114
FIJI . . . . .	114
Frequency Domain . . . . .	114
Gaussian Process . . . . .	114
Generative Adversarial Networks (GANs) . . . . .	114
Genetic Algorithms . . . . .	115
Ground Truth . . . . .	115
Hallucinations . . . . .	115
Hyperparameters . . . . .	115
Image Classification . . . . .	115
Image Restoration . . . . .	116
Instance Segmentation . . . . .	116
IoU . . . . .	116
Manual Annotation . . . . .	116
Metadata . . . . .	116
N2N (Noise2Noise) . . . . .	116
N2S (Noise2Self) . . . . .	117
N2V (Noise2Void) . . . . .	117
Nonlinear Problem . . . . .	117
Object Detection . . . . .	117

Panoptic Segmentation . . . . .	117
Pixel Classifiers . . . . .	117
Point Spread Function (PSF) . . . . .	118
Quality Control Metric . . . . .	118
RCAN (residual channel attention network) . . . . .	118
ReLU . . . . .	118
Self-supervised learning . . . . .	118
Semantic Segmentation . . . . .	118
Sigmoid Function . . . . .	119
Spatial Domain . . . . .	119
Star-convex Polygon . . . . .	119
Supervised learning . . . . .	119
Training Data . . . . .	119
Transfer Learning . . . . .	119
Transformer Models . . . . .	120
True Negatives . . . . .	120
True Positives . . . . .	120
Virtual Machine . . . . .	120
Zernike Modes . . . . .	121
<b>References</b>	<b>122</b>
<b>Appendices</b>	<b>135</b>
<b>A 3D-RCAN for Image Restoration</b>	<b>135</b>
A.1 Brief Introduction to the Model . . . . .	135
A.1.1 Residual Learning . . . . .	135
A.1.2 Channel Attention Mechanism . . . . .	136
A.1.3 3D Adaptation for Image Volumes . . . . .	136
A.1.4 Channel Reduction for Efficient 3D Processing . . . . .	136
A.2 Overview of the Code and Data . . . . .	138
A.2.1 System Requirements . . . . .	138
A.2.2 Installation of Dependencies . . . . .	138
A.2.3 Quick Setup Steps . . . . .	138
A.2.4 Dataset . . . . .	139
A.2.5 Training . . . . .	139
A.2.6 Applying the Model . . . . .	140
A.3 Denoising Tutorial . . . . .	141
A.3.1 Data Preparation . . . . .	141
A.3.2 Training a Denoising Model . . . . .	142
A.3.3 Applying the Denoising Model . . . . .	143
A.3.4 Further Denoising Guidance . . . . .	145

A.4	Deconvolution Tutorial . . . . .	146
A.4.1	Data Preparation . . . . .	146
A.4.2	Training a Deconvolution Model . . . . .	147
A.4.3	Applying the Deconvolution Model . . . . .	148
A.4.4	Further Deconvolution Guidance . . . . .	149
A.5	Aberration Correction Tutorial . . . . .	150
A.5.1	Data Preparation . . . . .	150
A.5.2	Training an Aberration Correction Model . . . . .	151
A.5.3	Applying the Aberration Correction Model . . . . .	152
A.5.4	Further Aberration Correction Guidance . . . . .	153
A.6	Resolution Enhancement Tutorial . . . . .	153
A.6.1	Data Preparation . . . . .	153
A.6.2	Training a Resolution Enhancement Model . . . . .	155
A.6.3	Applying the Resolution Enhancement Model . . . . .	156
A.6.4	Further Resolution Enhancement Guidance . . . . .	157

# Welcome

This is an initial outline for the welcome page:

- Include a very brief introduction of the book (this is not the introduction chapter).
- Explain how to interact with the book. For example, explain how code snippets work for the reader, links to figures, glossary terms, etc.
- State the licensing and use restrictions.
- How to cite the book.

# 1 Preface

AI at Every Stage of the Microscopy Workflow

## 1.1 Introduction

The rapid advancement of microscopy, further fueled by parallel development in other related technologies such as probe chemistry, molecular biology, image analysis methods, has propelled bioimaging to the forefront of life science studies Balasubramanian et al. [6]. In fact, many life scientists now fully expect to visualize intricate biological processes in the context of a whole, living organism – a prospect deemed impossible just a mere decade ago.

This seemingly limitless possibility has greatly amplified the complexity of the questions that can be probed. Likewise, it demands a significant degree of forethought in the design of biological experiments [112, 57] to ensure that the maximal amount of biological information can be extracted from the acquired data. On the other hand, the wealth of information contained in modern bioimage data can now often exceed the grasp of a human observer and thus require computational means to aid in data interpretation. In other words, the need for the ever-expanding and multifaceted field of artificial intelligence (AI) can no longer be separated from bioimaging. Yet the absence of a comprehensive resource to systematically explore how AI can enhance bioimaging prevents many in the community from taking advantage of this indispensable technology. This textbook therefore aims to demystify AI in its many forms and outline proper use cases thereof to facilitate biological discovery. We have specifically contextualized the textbook for our two target audiences: microscopists and biologists. The scientists developing new AI and ML tools often fail to tailor explanations and tutorials for these users who will ultimately facilitate biological discovery. We hope this textbook will bridge the gap.

## 1.2 Textbook Outline

As the chapters of this book will collectively demonstrate, AI now plays an indispensable role in the four main stages of the microscopy experimental workflow (Figure 1.1): Project

design, tool-building, execution of experiments, and data modeling. To address each part of this workflow, this textbook is divided into three distinct sections.

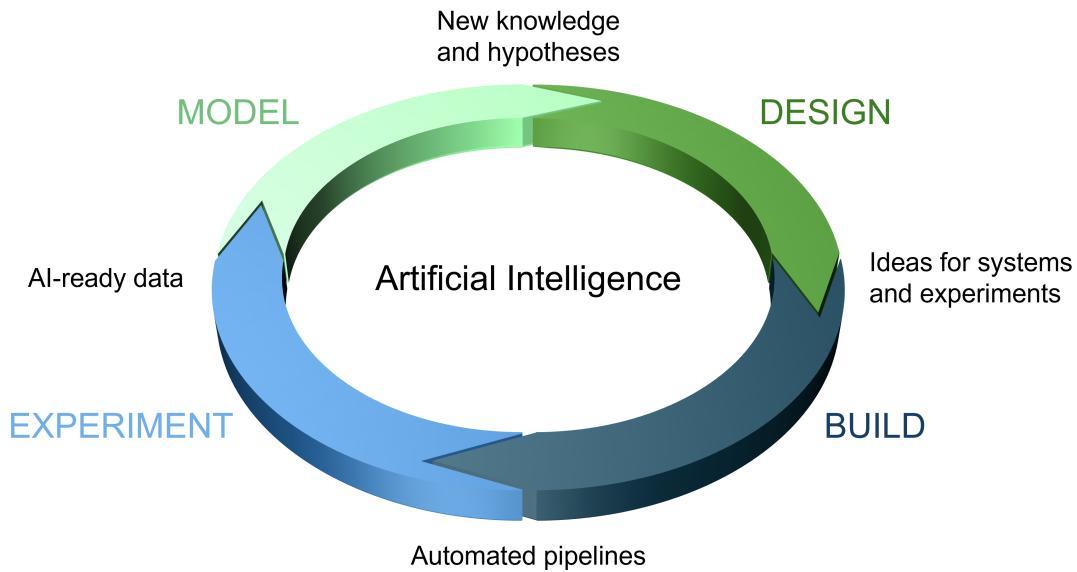


Figure 1.1: **AI impacts every aspect of the microscopy workflow.** The experimental process represents a cycle through four distinct phases: experimental design, constructing systems and workflows, testing hypotheses using these workflows, and analyzing their outputs to develop new models which inform new experiments. In the current age, AI holds the capacity to impact each phase of this cycle.

The first focuses on introducing the core concepts of AI, large language models (LLMs) and the computational architectures that underpin these technologies. [Chapter 2](#) will introduce the basic principles behind artificial intelligence and machine learning (ML) as a gateway for those less familiar with this topic. In the experimental workflow, modern hypotheses can now drive a project design so complex that artificial intelligence may be required to conceptualize and build the dedicated microscopy tool needed to tackle the experimental challenge. This can be accomplished by using large language models (LLMs) as assistants, as discussed in [Chapter 3](#). [Chapter 4](#) will further explain the construction of AI models and how the accuracy of their outputs can be validated.

The second section focuses on extending microscopy hardware with AI. To render a microscope “smart”, i.e., for it to perform an experiment with minimal or no human intervention, ML software first needs to be trained on extensive and high-quality training data. The collection and design of training data is covered in [Chapter 5](#). [Chapter 6](#) will discuss the implementation of AI and ML to enhance the capabilities of existing microscopes through resolution improvement and denoising. Likewise, AI and ML models can be designed to control event-driven image

acquisition to visualize biological events that are rare and/or photosensitive ([Chapter 7](#)).

Finally, the third section addresses the use of AI and ML to process and analyze microscopy data in order to extract biological information beyond what is achievable through human observation alone. [Chapter 8](#) will focus on helping readers find and use available open-access tools. When those existing tools are yet insufficient, [Chapter 9](#) seeks to further widen the horizon by discussing the process of training new models for specific image processing and analysis tasks. Finally, [Chapter 10](#) will address the question of “when is my model performing well enough?”. This important step pinpoints potential weaknesses of any model, thereby placing results in meaningful context.

We conclude with a [chapter](#) that summarizes the core concepts and hopes to inspire readers to begin taking full advantage of the promise of the new era of AI and ML. To close the loop, we also hope that by equipping readers with the [proper vocabulary](#), the book will facilitate more effective communication between biologists and AI tool developers.

# **Part I**

# **Getting Started with AI**

# 2 AI Primer

An Introduction to Artificial Intelligence

Under your first header, include a brief introduction to your chapter.

Starting prompt for this chapter: Chapter 2 demystifies Artificial Intelligence for microscopy users. It should define terms (e.g., machine/deep learning, supervised/unsupervised learning) without programming details such that an educated scientist without AI experience can understand how these concepts apply to microscopy in life sciences. The use-cases and strengths of different approaches for different applications should be discussed (e.g., contrasting unsupervised clustering vs supervised segmentation). This chapter should broadly introduce image restoration and segmentation, as they will be themes throughout.

Suggestion from authors' meetings: This chapter can draw on the outlines from other chapters to introduce key topics for the following chapters.

## 2.1 Include section headers as appropriate

Use markdown heading level two for section headers. You can use standard markdown formatting, for example *emphasize the end of this sentence*.

This is a new paragraph with more text. Your paragraphs can cross reference other items, such as Figure 11.1. Use `fig` to reference figures, and `eq` to reference equations, such as Equation 11.1.

### 2.1.1 Sub-subsection headers are also available

To make your sections cross reference-able throughout the book, include a section reference, as shown in the header for Section 11.4.

## 2.2 Bibliography and Citations

To cite a research article, add it to references.bib and then refer to the citation key. For example, reference Stringer et al. [107] refers to CellPose and reference Chamier et al. [16] refers to ZeroCostDL4Mic.

## 2.3 Adding to the Glossary

We are using R code to create a glossary for this book. To add a definition, edit the glossary.yml file. To reference the glossary, enclose the word as in these examples: LLMs suffer from **hallucinations**. It is important to understand the underlying **training data**, validation data and **false positives** to interpret your results. Clicking on the word will reveal its definition by taking you to the entry on the [Glossary](#) page. Pressing back in your browser will return you to your previous place in the textbook.

## 2.4 Code and Equations

This is an example of including a python snippet that generates a figure

```
import matplotlib.pyplot as plt
plt.plot([1,23,2,4])
plt.show()
```

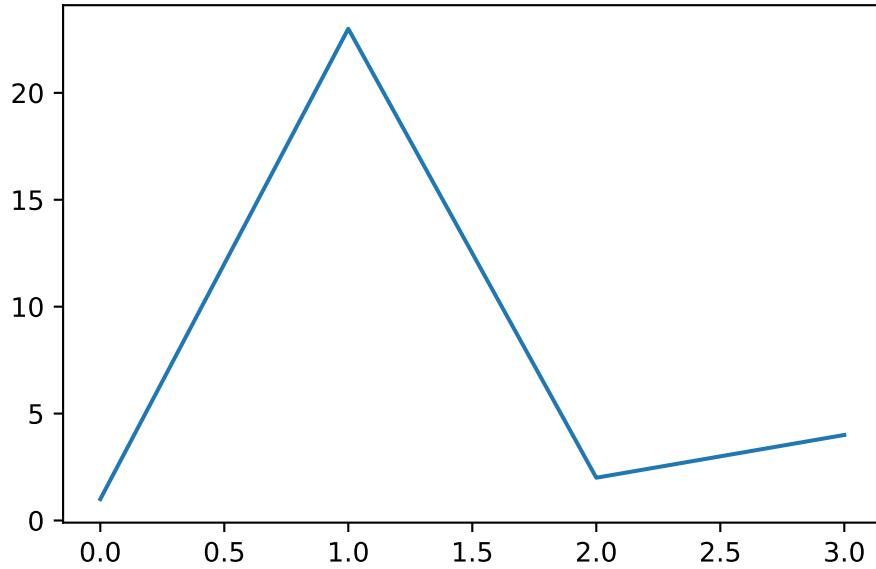


Figure 2.1: Simple Plot

In some cases, you may want to include a code-block that is not executed when the book is compiled. Use the `eval: false` option for this.

```
import matplotlib.pyplot as plt
plt.plot([1,23,2,4])
plt.show()
```

Figures can also be generated that do not show the code by using the option for `code-fold: true`.

```
import numpy as np
import matplotlib.pyplot as plt

r = np.arange(0, 2, 0.01)
theta = 2 * np.pi * r
fig, ax = plt.subplots(
    subplot_kw = {'projection': 'polar'}
)
ax.plot(theta, r)
ax.set_rticks([0.5, 1, 1.5, 2])
ax.grid(True)
plt.show()
```

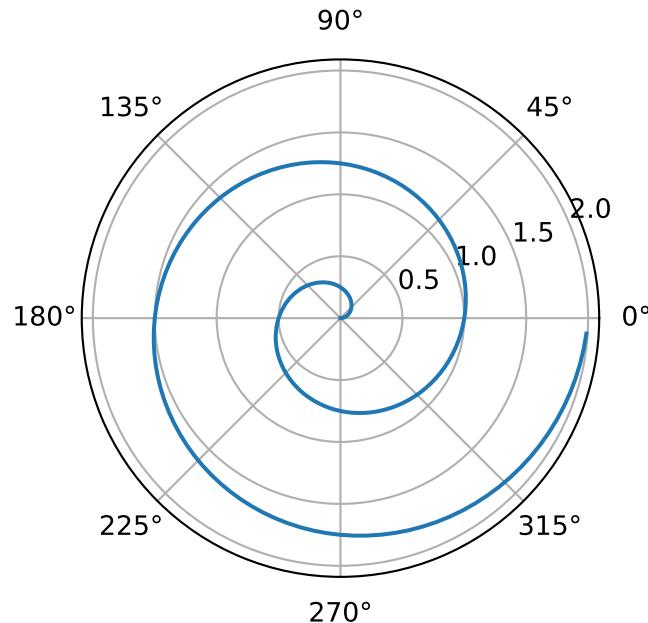


Figure 2.2: A spiral on a polar axis

Here is an example equation.

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (2.1)$$

#### 2.4.1 Embedding Figures

You can also embed figures from other notebooks in the repo as shown in the following embed example.

When embedding notebooks, please store the .ipynb file in the notebook directory. Include the chapter in the name of your file. For example, `chapter4_example_u-net.ipynb`. This is how we will handle chapter- or example-specific environments. We will host notebooks on Google Colab so that any required packages for the code—but not for rendering the book at large—will be installed there. That way, we will not need to handle a global environment across the book.

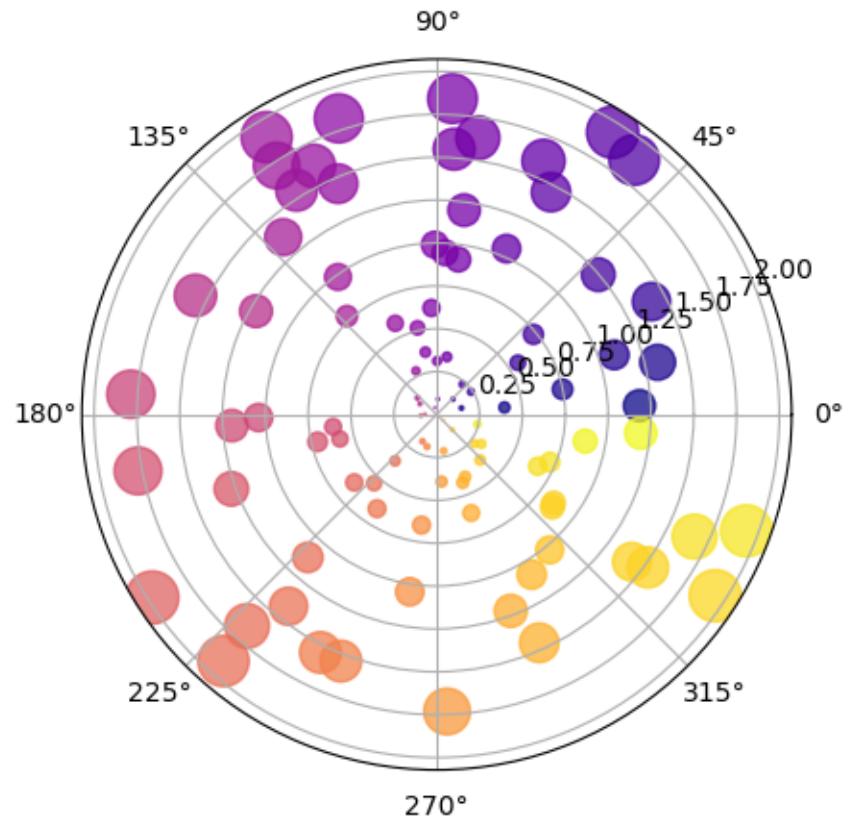


Figure 2.3: Polar plot of circles of random areas at random coords

## 2.5 Quarto has additional features.

You can learn more about markdown options and additional Quarto features in the [Quarto documentation](#). One example that you might find interesting is the option to include callouts in your text. These callouts can be used to highlight potential pitfalls or provide additional optional exercises that the reader might find helpful. Below are examples of the types of callouts available in Quarto.

### Note

Note that there are five types of callouts, including: `note`, `tip`, `warning`, `caution`, and `important`. They can default to open (like this example) or collapsed (example below).

### Tip

These could be good for extra material or exercises.

### Caution

There are caveats when applying these tools. Expand the code below to learn more.

```
r = np.arange(0, 2, 0.01)
theta = 2 * np.pi * r
```

### Warning

Be careful to avoid hallucinations.

### Important

This is key information.

# 3 Large Language Models (LLMs)

Large Language Models and AI Agents for Microscopy Imaging

In recent years, large language models (LLMs) have revolutionized how we interact with technology, bringing unprecedented capabilities to scientific research including microscopy. This chapter explores how microscopists can leverage these powerful AI tools to enhance their workflow, from learning concepts to automating analysis tasks. We'll discuss both general-purpose and microscopy-specific tools while highlighting practical applications and potential pitfalls.

## 3.1 Foundations of Large Language Models

This section introduces the fundamental concepts behind modern language models, focusing on transformer architectures that power tools like ChatGPT. We'll explain how these models function, their capabilities for understanding scientific text, and their emerging role in generating code for image analysis tasks. We'll demonstrate how microscopists can effectively use LLMs to learn new concepts, troubleshoot methods, and generate starting points for analysis scripts.

## 3.2 Multi-modal AI: Vision-Language Models and Generative AI

Moving beyond text-only interfaces, multi-modal models combine language understanding with visual processing capabilities. This section explores how Vision-Language Models (VLMs) like GPT-4o can “see” and interpret microscopy images, assist with image annotation, and even aid in experimental design. We'll also cover generative AI technologies including diffusion models that can create synthetic training data, perform style transfer, or convert microscopy images into vector graphics for publications.

### **3.3 AI Agents for Microscopy Workflows**

AI agents represent the next evolution - autonomous systems that combine language understanding with specialized scientific knowledge and the ability to execute actions. We'll examine microscopy-specific tools like Omega and the BioImage.io chatbot that can perform complex bioimage analysis workflows through natural language instructions. This section will explore chain-of-thought reasoning, code generation and execution capabilities, and how these agents use visual feedback to iteratively improve results.

### **3.4 Challenges and Limitations**

While powerful, AI assistants come with significant limitations that microscopists must understand. This section addresses critical challenges including:

- and factual errors in generated content
- The “black box” nature of models and concerns about reproducibility
- Alignment problems when tools lack domain-specific knowledge
- The need for human validation and the dangers of overreliance
- Practical strategies for steering models toward scientifically valid outputs

### **3.5 Future Directions**

The intersection of LLMs and microscopy is rapidly evolving. This final section examines emerging capabilities and future possibilities, including:

- Generalist vision-language models capable of performing diverse analysis tasks
- Models that can directly transform input images into processed outputs
- The integration of AI agents with microscope hardware for fully autonomous imaging
- Smart microscopy systems that adapt acquisition parameters based on real-time image understanding
- Ethical considerations and best practices for responsible AI adoption in biological research

### **3.6 Practical Guide: Getting Started with LLMs for Microscopy**

This hands-on section provides step-by-step guidance for microscopists to begin leveraging LLMs effectively, including:

- Crafting effective prompts that produce reliable, scientific outputs
- Using ChatGPT and similar tools to learn imaging concepts and generate analysis code
- Getting started with BioImage.io tools and microscopy-specific AI agents
- Strategies for validating and verifying AI-generated solutions
- Example workflows demonstrating LLM integration into real microscopy analysis tasks

# 4 Architectures and Loss Models

optionally add a subtitle

Under your first header, include a brief introduction to your chapter.

Starting prompt for this chapter: Chapter 4 introduces architectures and loss models, defining them and providing examples through two practical case studies: image restoration and segmentation. Although this chapter will include code snippets/exercises, the presentation of essential concepts should communicate the philosophy behind the choice of a model for non-programmers.

## 4.1 Include section headers as appropriate

Use markdown heading level two for section headers. You can use standard markdown formatting, for example *emphasize the end of this sentence*.

This is a new paragraph with more text. Your paragraphs can cross reference other items, such as Figure 11.1. Use `fig` to reference figures, and `eq` to reference equations, such as Equation 11.1.

### 4.1.1 Sub-subsection headers are also available

To make your sections cross reference-able throughout the book, include a section reference, as shown in the header for Section 11.4.

## 4.2 Bibliography and Citations

To cite a research article, add it to references.bib and then refer to the citation key. For example, reference Stringer et al. [107] refers to CellPose and reference Chamier et al. [16] refers to ZeroCostDL4Mic.

## 4.3 Adding to the Glossary

We are using R code to create a glossary for this book. To add a definition, edit the glossary.yml file. To reference the glossary, enclose the word as in these examples: LLMs suffer from [hallucinations](#). It is important to understand the underlying [training data](#), validation data and [false positives](#) to interpret your results. Clicking on the word will reveal its definition by taking you to the entry on the [Glossary](#) page. Pressing back in your browser will return you to your previous place in the textbook.

## 4.4 Code and Equations

This is an example of including a python snippet that generates a figure

```
import matplotlib.pyplot as plt
plt.plot([1,23,2,4])
plt.show()
```

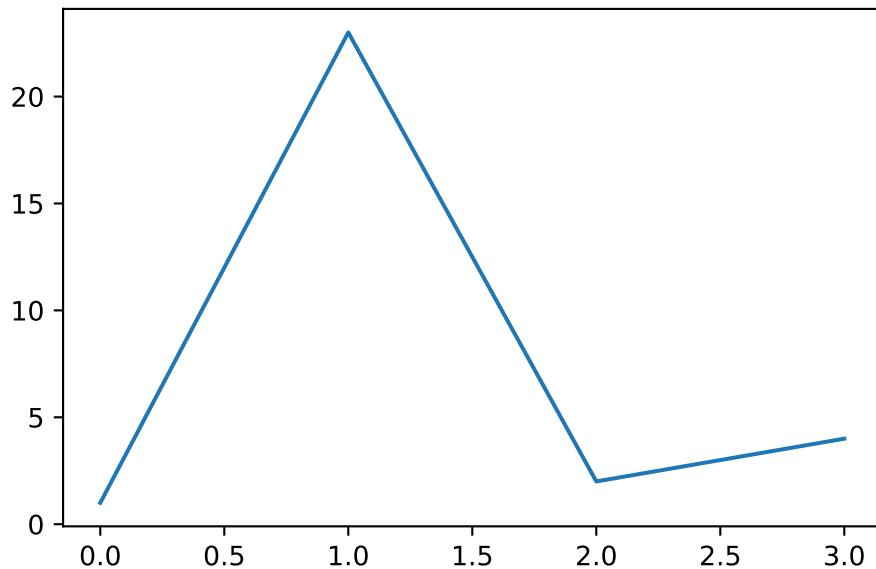


Figure 4.1: Simple Plot

In some cases, you may want to include a code-block that is not executed when the book is compiled. Use the `eval: false` option for this.

```
import matplotlib.pyplot as plt
plt.plot([1,23,2,4])
plt.show()
```

Figures can also be generated that do not show the code by using the option for `code-fold: true`.

```
import numpy as np
import matplotlib.pyplot as plt

r = np.arange(0, 2, 0.01)
theta = 2 * np.pi * r
fig, ax = plt.subplots(
    subplot_kw = {'projection': 'polar'}
)
ax.plot(theta, r)
ax.set_rticks([0.5, 1, 1.5, 2])
ax.grid(True)
plt.show()
```

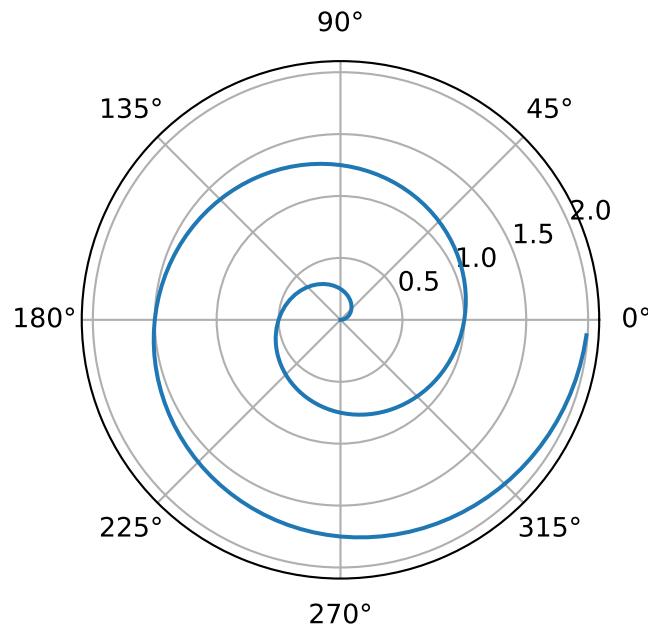


Figure 4.2: A spiral on a polar axis

Here is an example equation.

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (4.1)$$

#### 4.4.1 Embedding Figures

You can also embed figures from other notebooks in the repo as shown in the following embed example.

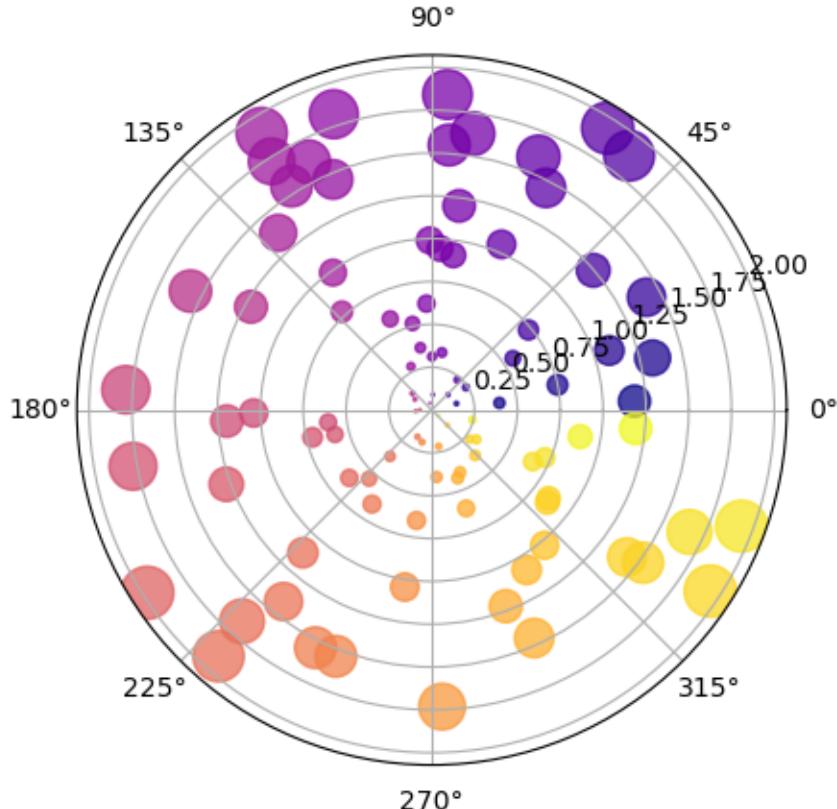


Figure 4.3: Polar plot of circles of random areas at random coords

When embedding notebooks, please store the .ipynb file in the notebook directory. Include the chapter in the name of your file. For example, `chapter4_example_u-net.ipynb`. This is how we will handle chapter- or example-specific environments. We will host notebooks on Google Colab so that any required packages for the code—but not for rendering the book at large—will be installed there. That way, we will not need to handle a global environment across the book.

## 4.5 Quarto has additional features.

You can learn more about markdown options and additional Quarto features in the [Quarto documentation](#). One example that you might find interesting is the option to include callouts in your text. These callouts can be used to highlight potential pitfalls or provide additional optional exercises that the reader might find helpful. Below are examples of the types of callouts available in Quarto.

### Note

Note that there are five types of callouts, including: `note`, `tip`, `warning`, `caution`, and `important`. They can default to open (like this example) or collapsed (example below).

### Tip

These could be good for extra material or exercises.

### Caution

There are caveats when applying these tools. Expand the code below to learn more.

```
r = np.arange(0, 2, 0.01)
theta = 2 * np.pi * r
```

### Warning

Be careful to avoid hallucinations.

### Important

This is key information.

## **Part II**

# **Image Acquisition**

# 5 Collecting Training Data

## Image Collection and Considerations

Under your first header, include a brief introduction to your chapter.

Starting prompt for this chapter: Chapter 5 discusses collecting, annotating and validating training data. It should highlight potential pitfalls such as balanced data sets, out-of-distribution problems, etc. It should also address the question: how do you collect training data on your microscope? For example, this chapter should discuss collecting low/high-laser power pairs for the purpose of training an image restoration model.

### 5.1 Include section headers as appropriate

Use markdown heading level two for section headers. You can use standard markdown formatting, for example *emphasize the end of this sentence*.

This is a new paragraph with more text. Your paragraphs can cross reference other items, such as Figure 11.1. Use `fig` to reference figures, and `eq` to reference equations, such as Equation 11.1.

#### 5.1.1 Sub-subsection headers are also available

To make your sections cross reference-able throughout the book, include a section reference, as shown in the header for Section 11.4.

### 5.2 Bibliography and Citations

To cite a research article, add it to references.bib and then refer to the citation key. For example, reference Stringer et al. [107] refers to CellPose and reference Chamier et al. [16] refers to ZeroCostDL4Mic.

## 5.3 Adding to the Glossary

We are using R code to create a glossary for this book. To add a definition, edit the glossary.yml file. To reference the glossary, enclose the word as in these examples: LLMs suffer from [hallucinations](#). It is important to understand the underlying [training data](#), validation data and [false positives](#) to interpret your results. Clicking on the word will reveal its definition by taking you to the entry on the [Glossary](#) page. Pressing back in your browser will return you to your previous place in the textbook.

## 5.4 Code and Equations

This is an example of including a python snippet that generates a figure

```
import matplotlib.pyplot as plt
plt.plot([1,23,2,4])
plt.show()
```

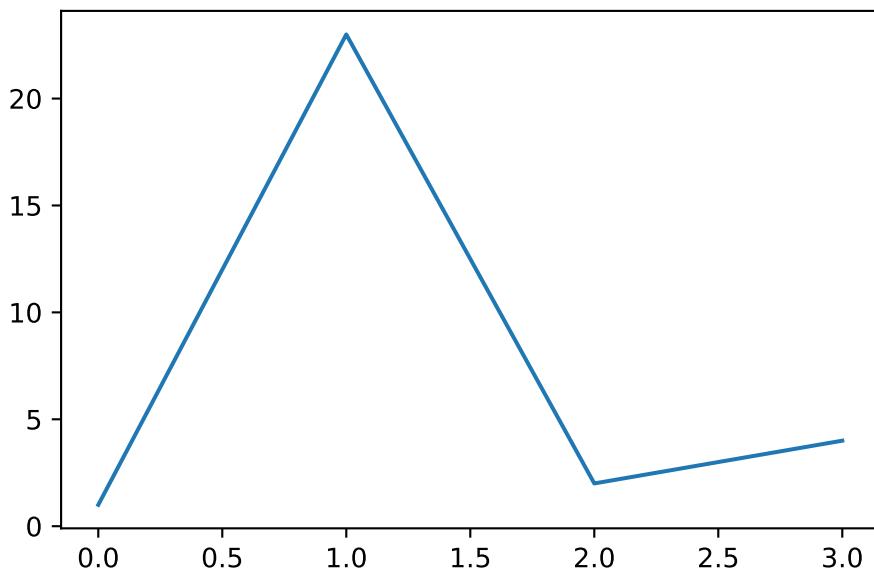


Figure 5.1: Simple Plot

In some cases, you may want to include a code-block that is not executed when the book is compiled. Use the `eval: false` option for this.

```
import matplotlib.pyplot as plt
plt.plot([1,23,2,4])
plt.show()
```

Figures can also be generated that do not show the code by using the option for `code-fold: true`.

```
import numpy as np
import matplotlib.pyplot as plt

r = np.arange(0, 2, 0.01)
theta = 2 * np.pi * r
fig, ax = plt.subplots(
    subplot_kw = {'projection': 'polar'}
)
ax.plot(theta, r)
ax.set_rticks([0.5, 1, 1.5, 2])
ax.grid(True)
plt.show()
```

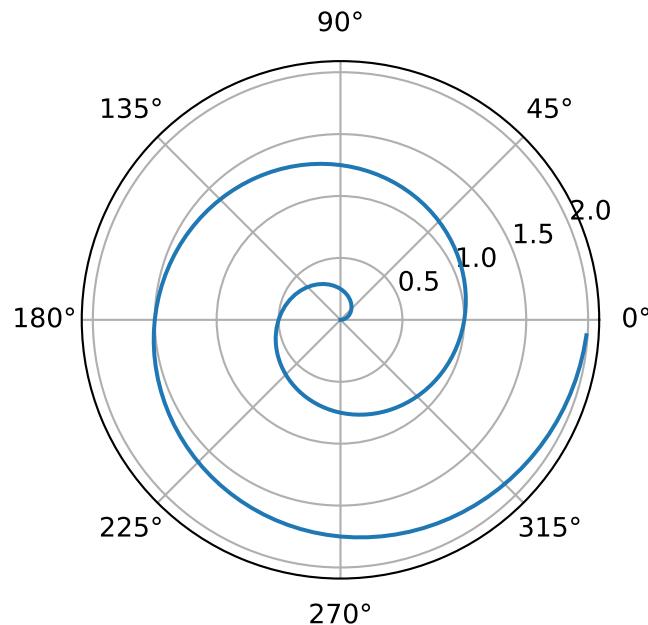


Figure 5.2: A spiral on a polar axis

Here is an example equation.

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (5.1)$$

#### 5.4.1 Embedding Figures

You can also embed figures from other notebooks in the repo as shown in the following embed example.

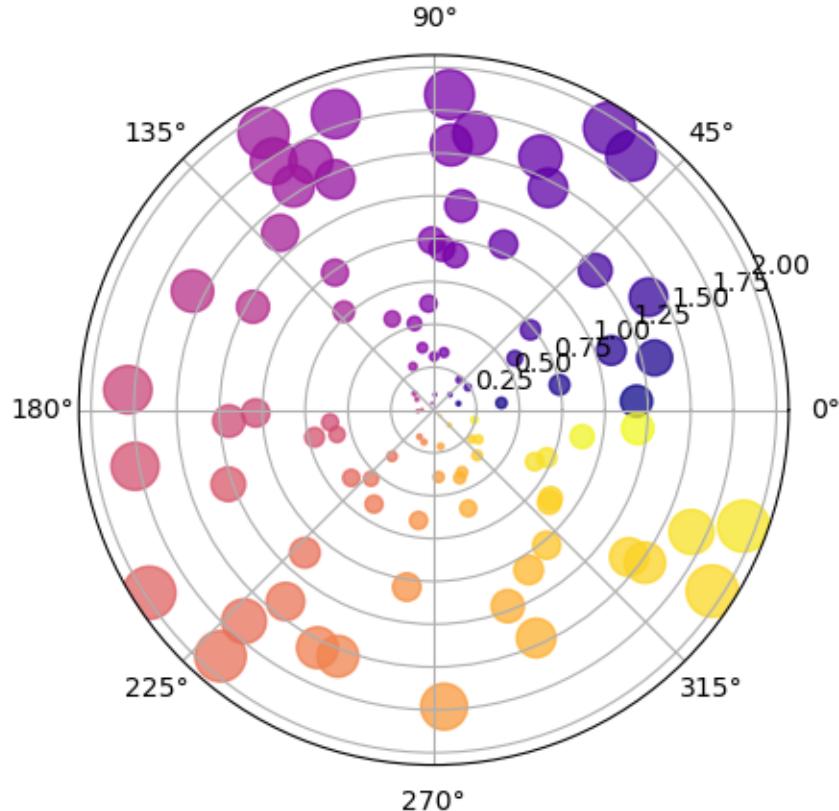


Figure 5.3: Polar plot of circles of random areas at random coords

When embedding notebooks, please store the .ipynb file in the notebook directory. Include the chapter in the name of your file. For example, `chapter4_example_u-net.ipynb`. This is how we will handle chapter- or example-specific environments. We will host notebooks on Google Colab so that any required packages for the code—but not for rendering the book at large—will be installed there. That way, we will not need to handle a global environment across the book.

## 5.5 Quarto has additional features.

You can learn more about markdown options and additional Quarto features in the [Quarto documentation](#). One example that you might find interesting is the option to include callouts in your text. These callouts can be used to highlight potential pitfalls or provide additional optional exercises that the reader might find helpful. Below are examples of the types of callouts available in Quarto.

### Note

Note that there are five types of callouts, including: `note`, `tip`, `warning`, `caution`, and `important`. They can default to open (like this example) or collapsed (example below).

### Tip

These could be good for extra material or exercises.

### Caution

There are caveats when applying these tools. Expand the code below to learn more.

```
r = np.arange(0, 2, 0.01)
theta = 2 * np.pi * r
```

### Warning

Be careful to avoid hallucinations.

### Important

This is key information.

## 6 Image Restoration

### Using Artificial Intelligence for Image Restoration in Fluorescence Microscopy

Fluorescence microscopy is central to biological discovery, enabling the visualization of cellular and subcellular structures with exquisite detail. From observing dynamic intracellular processes to mapping entire tissues, microscopy is indispensable for understanding biological systems Wu and Shroff [120]. However, the quality of microscopic images is often compromised due to intrinsic limitations such as noise, optical aberrations, diffraction, and limited signal Schermelleh, Heintzmann, and Leonhardt [96]. These factors hinder analysis and interpretation, particularly when studying fine biological structures or dynamic processes. For example, by quantifying the intensity of fluorescence signals, researchers infer the abundance or expression levels of specific molecules (e.g. proteins tagged with a genetically expressed marker). Using image segmentation, researchers can analyze the size, shape, distribution and number of specific objects within a defined region Archit et al. [3]. By tracking the movement of fluorescently labeled molecules or cells over time, researchers can study processes including endocytosis, intracellular transport, and cellular signaling Sahl, Hell, and Jakobs [93]. In these scenarios, noise and low SNR can make it difficult to distinguish between the actual signal and background; and spatial blurring causes fluorescence signal to spread out, confounding the ability to accurately assign pixels to specific regions or structures.

Many of these problems can be overcome with suitable hardware or by using advanced imaging techniques. To improve spatial resolution, super-resolution techniques (stimulated emission depletion (STED), structured illumination microscopy (SIM), or single-molecule localization microscopy (SMLM)) may be employed Schermelleh et al. [97]. To suppress noise and maximize the collection of useful signal, highly sensitive detectors, such as cooled charge-coupled devices (CCDs) or complementary metal-oxide-semiconductor (CMOS) sensors can be used. To correct optical distortions, advanced microscopes integrate adaptive optics (AO) [46, 38], which use real-time feedback to dynamically adjust the focus and compensate for aberrations. However, these technologies often come with high cost, require complex operation, or need extensive maintenance. Additionally, improving one attribute of the image (e.g. spatial resolution) often results in a compromise in another (e.g., temporal resolution), requiring careful consideration of the specific needs of a study and the resources available Shroff et al. [104].

Another way to address these limitations is to develop [image restoration](#) techniques that enhance the quality of microscopic images, often without expensive instrumentation. Tradition-

ally, these techniques relied on mathematical algorithms and physical models of the imaging process. However, recent advances in artificial intelligence (AI), particularly deep learning, have transformed image restoration by enabling data-driven approaches that offer improved performance and flexibility. Manufacturers have increasingly integrated image restoration technologies as an essential component of their products to ensure their users can obtain clearer images directly from the microscope, e.g., Leica THUNDER Imager, Olympus cellSens, Andor iQ, ZEISS arivis, and Nikon NIS-Elements.

In this chapter, we provide a practical overview of image restoration for applications in fluorescence microscopy. We begin by introducing the general concept of image restoration, then describe traditional and deep learning-based approaches. Next, we review key technologies and advances by categorizing restoration into four major areas: denoising, [deconvolution](#), [deaberration](#), and resolution enhancement. Finally, we provide practical guidelines for implementing image restoration, including step-by-step workflows and test datasets for interested readers to practice applying these methods.

## 6.1 General Concepts in Image Restoration

Image restoration is the application of mathematical and computational techniques aimed at improving the quality of an image by reversing or reducing the effects of degradation that may occur during the imaging process. These degradations stem from noise, blur, aberrations, and other artifacts. The goal of the restoration is to recover or estimate the latent (true) image from the degraded raw data. Image restoration has the potential not only to improve the precision of biological analyses but also to expand the capabilities of existing microscopes, enabling researchers using basic hardware to achieve advanced imaging quality.

### 6.1.1 Image Degradation Model

Understanding the types of distortions and the causes behind them is the first step in image restoration. This could involve a mathematical model that characterizes how an image is corrupted, using the concept that any real-world imaging system cannot capture an object perfectly. Instead, the imaging system or sample introduces imperfections, which distort the image. In fluorescence microscopy, a general image degradation model can be written as:

$$i = o \otimes f(\phi) + n \quad (6.1)$$

Where  $i$  is the image acquired by the microscope;  $o$  is the intensity distribution of the biological sample;  $f(\phi)$  is the point spread function ([PSF](#)) of the system;  $\phi$  is the phase aberration or wavefront distortion (when  $\phi = 0$ , the wavefront is ‘flat’ or aberration-free and  $f$  is the ideal PSF predicted by theory);  $\otimes$  is the [convolution](#) operator; and  $n$  models noise contamination. Two common types of noise are Gaussian noise and Poisson noise. Gaussian noise is caused

by random fluctuations in the image due to imperfections in the camera sensor or electronic interference. Poisson noise arises from the random nature of photon detection, meaning that the number of photons detected can vary from one measurement to the next, especially at low light levels. Equation 6.1 is described schematically in Figure 6.1.

### 6.1.2 Image Restoration with Traditional Approaches

Traditional microscopy image restoration techniques use mathematical modeling of the imaging model and statistical analyses that attempt to reverse or mitigate image degradations, and/or additional microscope hardware. We find it helpful to categorize image restoration into four main areas: denoising, deconvolution, deaberration, and resolution enhancement.

#### 6.1.2.1 Denoising

Traditional denoising methods try to remove noise while preserving important features such as edges and fine structures. Many denoising algorithms assume Gaussian noise, for computational tractability. Gaussian noise removal can be achieved by filtering-based methods in the [spatial](#) or [frequency domain](#), such as Gaussian filtering, mean filtering, wavelet filtering, bilateral filtering Venkatesh, Mohan, and Selamantula [110], and non-local-based BM3D Danielyan et al. [25]. Simple linear filters are easy to implement, but they cause a loss in high-frequency information. Complex denoising methods require careful parameter design and are usually computationally intensive. In many cases, the Gaussian noise model provides a good approximation, but Poisson noise is also a key source of noise for fluorescence microscopy given the quantized nature of fluorescence emission, especially under low signal conditions when detector noise is minimal. One method for dealing with Poisson noise Zhang et al. [124] directly incorporates its statistics, e.g. the PURE-LET method Li, Luisier, and Blu [59]. Alternatively, a nonlinear variance-stabilizing transformation (VST) Makitalo and Foi [68] can be used to convert the Poisson denoising problem into a Gaussian denoising problem. There are several free and open-source resources available for microscopy image denoising, including filters and the PureDeNoise Luisier et al. [67] plugin in ImageJ/[Fiji](#), and the scikit-image library in Python Walt et al. [113].

#### 6.1.2.2 Deconvolution

Deconvolution is the process of reversing optical blur introduced by the PSF of the microscope ( $f(\phi)$  in Figure 6.1) and is most often implemented without considering optical aberrations. It improves effective contrast and resolution by accounting for such blur and reassigning the relevant signal to its most likely location, given the blurring and noise models. Traditional deconvolution methods include frequency-domain algorithms like naive inverse filtering and Wiener filtering[118, 32], optimization-based algorithms like Tikhonov regularization Tikhonov [108], and iterative deconvolution methods like the Tikhonov-Miller algorithm Miller [69], fast

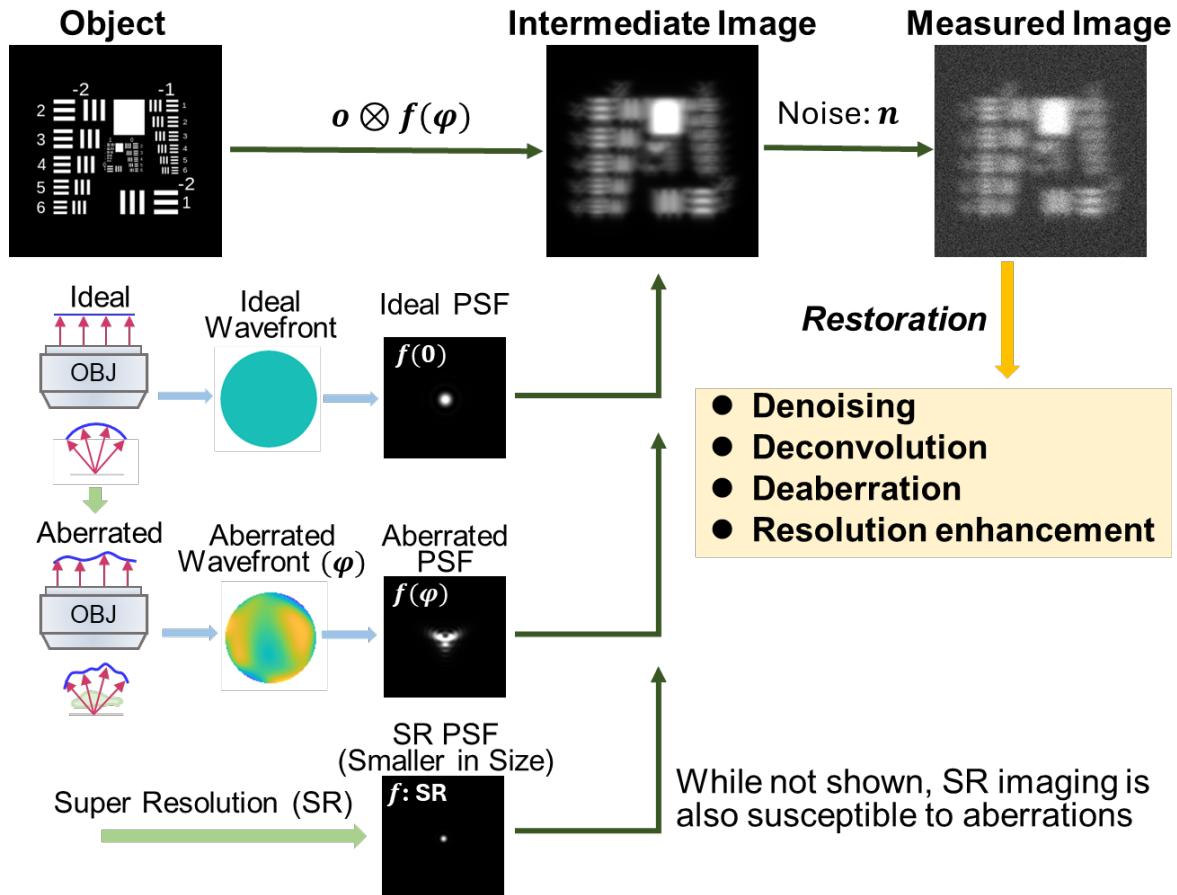


Figure 6.1: **Schematic of image degradation model for fluorescence microscopy.** Object information passes through the imaging system with degradation due to blurring from the point spread function (PSF, here denoted by the convolution kernel  $f$ ) and noise. In an ideal case, there is no distortion, and the wavefront associated with a point source maintains a flat phase at the back focal plane of the microscope objective. In any real situation, wavefront distortion due to optical imperfections or heterogeneity within the biological sample results in an aberrated wavefront (indicated by  $\varphi$ ) and PSF, which results in additional blurring. If spatial resolution is enhanced, e.g., by super-resolution methods (SR), the PSF becomes smaller, resulting in sharper images. Super-resolution imaging is also (typically more) prone to aberrations than conventional diffraction-limited imaging, and noise is inherent to any imaging process.

iterative soft-thresholding algorithm Beck and Teboulle [10], and Richardson-Lucy deconvolution [66, 87]. Traditional deconvolution methods can require significant computational resources, especially for large datasets with complex blurring functions, and also can amplify noise [95, 33, 35]. Commercial deconvolution software includes Huygens, DeltaVison Deconvolution, and AutoQuant. There are also open-source deconvolution plugins integrated into Fiji Schindelin et al. [98] (e.g., DeconvolutionLab2 Sage et al. [91]) or MIPAV Bazin et al. [9] (Medical Image Processing, Analysis, and Visualization, <https://mipav.cit.nih.gov/>) programs.

#### 6.1.2.3 Deaberration

Aberrations refer to the optical imperfections or distortions that occur during image acquisition, degrading the quality of the captured image. Such aberrations can arise due to optical path length differences introduced anywhere in the imaging path, including instrument misalignment, optical imperfections, or differences in refractive index between the heterogeneous and refractile sample, immersion media, and/or objective immersion oil. Aberrations significantly alter the wavefront of light. When the wavefront is distorted, the light rays that are focused by the microscope do not converge as they ideally should. This distortion can lead to various image quality issues, such as loss of resolution, and distortion of fine features. Adaptive optics [46, 13] can mitigate aberrations by measuring wavefront distortions and subsequently compensating for them using a deformable mirror or other optical elements. However, implementing AO is nontrivial, often requiring additional control algorithms and new hardware, adding considerable expense to the underlying microscope.

#### 6.1.2.4 Resolution Enhancement

Optical super-resolution techniques, like STED, PALM, SIM, etc., bypass the diffraction limit by utilizing on/off fluorophore state transitions Hell [40], sophisticated hardware designs, and/or image reconstruction algorithms [93, 111, 119]. Alternatively, physical expansion of the sample can be used with conventional microscopes [18, 116]. While effective, all these methods present some tradeoff for the gain in spatial resolution, including increased acquisition time, additional illumination dose, specially designed probes, or more complex instrumentation [109, 19].

#### 6.1.2.5 Drawbacks

Although traditional image restoration algorithms are used extensively in all these categories, they suffer several drawbacks. For example, many methods require careful and manual parameter tuning, and they often perform poorly in challenging conditions, such as in the presence of high noise, low contrast, defocus, or complex background. This is often because they are based on idealized assumptions (Equation 6.1), which are often not met in practice. Another limitation of traditional restoration methods is that they are generally ‘content unaware’ and do not use sample-specific prior information (e.g., shape, size, intensity distributions). On the one hand,

this means that traditional algorithms generalize well, but on the other hand, they are not as performant as newer deep learning methods. Deep learning methods can learn a task such as denoising from the data themselves or provide a sample-specific prior Hagen et al. [37], and thus can outperform traditional methods in many cases.

### 6.1.3 Image Restoration with Deep Learning-Based Approaches

Artificial intelligence (AI), particularly deep learning, has achieved remarkable progress in recent years, leading to breakthroughs in image restoration tasks. Deep learning uses artificial neural networks with multiple layers to model and learn complex patterns in data through end-to-end training, eliminating the need for handcrafted features or manual parameter tuning. By leveraging large datasets and computational power, deep learning can capture non-linear relationships and subtle details within image data, making it particularly well-suited for image restoration. Techniques such as convolutional neural networks (CNNs), generative adversarial networks (GANs), [autoencoders](#), and [transfer learning](#) have been shown to tackle applications in denoising, deblurring, super-resolution, and deaberration. Recent advances, such as content-aware restoration ([CARE](#)) Weigert et al. [117], residual channel attention networks ([RCANs](#)) Chen et al. [21], and deep Fourier-based models Qiao et al. [84], have demonstrated improvements in image quality while reducing phototoxicity and photobleaching during acquisition. These models offer several advantages over more traditional methods of image restoration:

- **Automatic Feature Learning:** Traditional image restoration methods rely heavily on manually designed features (e.g., regularization selection) and parameter tuning, which often require fine adjustments for different types of images. Deep learning models, on the other hand, can automatically learn features from large datasets, eliminating the need for manual intervention. This significantly reduces the complexity of parameter tuning and makes the models more adaptable to complex samples or tasks.
- **Strong Non-Linear Modeling Capabilities:** Traditional algorithms are often based on linear image degradation models (Equation 6.1), which ignore sample-specific information and perform poorly in the presence of complex image distortions or noise. However, recovering the object structure from the acquired image is an inherently ill-posed and [nonlinear problem](#) due to noise and blurring. Deep learning uses nonlinear activation functions ([ReLU](#), [sigmoid](#)) and hierarchical layers to approximate complex relationships in the image data. Thus, deep learning can better model the global context of an image, allowing for more accurate restoration of image details under challenging and suboptimal imaging conditions. This can extend the use of existing hardware beyond its original use to new biological questions that were previously inaccessible.
- **Efficient Handling of Complex Scenes:** Deep learning models are typically trained on large datasets and are capable of handling complex scenes (e.g., images contaminated with noise and aberrations Hou et al. [42]) and large-scale data (e.g., GB- or TB-scale

time-lapse data). Once a model is trained, it can automatically adapt to new input data with similar acquisition parameters, which makes it much more efficient for large-scale image restoration tasks.

As deep learning models continue to evolve and integrate with advanced microscopy hardware and large datasets, they are expected to further push the boundaries of biological imaging and discovery. In the next section, we will introduce the application of deep learning methods in different image restoration tasks.

## 6.2 Deep Learning-Based Techniques for Image Restoration

Following the convention from Section 6.1.2, we group AI-based restoration into four broad categories.

### 6.2.1 Denoising

Traditional denoising techniques struggle to preserve fine details while removing noise [84, 123, 24]. Deep learning-based methods, on the other hand, can learn to predict what fine details look like even in the presence of noise. Such methods can be generally divided into **supervised learning** and **self-supervised learning** (Figure 6.2). Supervised learning methods require a large number of paired datasets of noisy and clean images for training. These datasets can be obtained by tuning illumination intensity and/or exposure time during image acquisition: when imaging with high intensity or longer exposure time, high-SNR clean images can be collected; otherwise, noisy images are collected. Training a supervised deep learning method on these data results in a mapping between noisy and clean images. When sufficient training data is available, these methods can provide very high-quality restoration on noisy data unseen by the network in the training process. However, paired datasets can be difficult to obtain in some microscopy settings (e.g., live-cell imaging). Typical denoising networks include U-Net-based Content-aware image restoration (CARE) Weigert et al. [117] and attention-based residual channel attention networks (RCAN) Chen et al. [21].

Self-supervised learning methods do not require paired noisy and clean images. Instead, only noisy images are needed for training. They exploit the inherent structure of the data, learning to predict parts of the image using information gleaned from other parts. Typical self-supervised methods include Noise2Void (**N2V**) Krull, Buchholz, and Jug [54] and Noise2Self (**N2S**) Batson and Royer [8]. The key idea underlying N2V and N2S is that pixels from clean images are often highly correlated, i.e., nearby pixels usually look similar and follow predictable patterns. For example, neighboring pixels could show smooth gradients in image texture. However, many noise sources are randomly correlated in space, particularly across large areas. So by selecting a suitable mechanism, noise and signal can be separated properly. In N2V, missing pixels are randomly masked in the noisy input image and the network is then trained to predict

the value of the missing pixel using the surrounding noisy pixels. In N2S, instead of masking the noisy image entirely, local patches are used, and the network is trained to predict the value of the noisy pixel from its neighbors. While self-supervised methods can be effective, they often do not perform as well as supervised methods when the sample or noise distribution is complex. These methods nevertheless are quite useful when paired noisy-clean images are difficult or impossible to obtain.

Noise2Noise (N2N) Lehtinen et al. [58] represents another kind of supervised training. It trains the model using two noisy images (A and B), where both images contain the same underlying clean signal. Essentially, it learns to map noisy image A to noisy image B. Since both images share the same clean signal beneath the noise, the network can learn to recover the underlying clean image. This method still relies on paired data as supervision, but the supervision signal comes from another noisy image rather than a traditional clean image.

### 6.2.2 Deconvolution

Deconvolution has long relied on iterative algorithms like the Richardson-Lucy method. Recent advances in deep learning (DL) have revolutionized this task, offering high speed and in some cases even more accurate predictions than traditional methods Li et al. [60]. Deep learning deconvolution methods can be divided into two categories: purely data-driven learning and physics-informed learning (Figure 6.3).

In the data-driven approach, training is conducted similarly to supervised denoising methods. The raw images are the low-quality blurred acquisitions; the high-quality reference can be the results of traditional restoration (e.g., multi-view jointly deconvolved results, or high SNR deconvolved results). Models like CARE, RCAN, and DenseDeconNet Guo et al. [35] can be used for deconvolution. Once trained, such data-driven deconvolution generally enables more rapid deconvolution than traditional methods. However, these methods are content-aware and depend heavily on high-quality paired training data. This criterion may introduce artifacts and affect how well the network generalizes.

Physics-informed methods integrate domain knowledge—such as the microscope’s point spread function (PSF), and image formation models—into the network architecture or loss functions. For example, the Richardson-Lucy Network (RLN) Li et al. [60] embeds Richardson-Lucy iterations within a convolutional network, creating a hybrid method that leverages classical knowledge and data-driven deep learning. MultiWienerNet uses multiple differentiable Wiener filters paired with a convolutional neural network, exploiting the knowledge of the system’s spatially varying PSF to quickly perform 2D and 3D reconstruction, achieving good results on spatially varying deconvolution tasks Yanny et al. [121]. Physics-informed methods can also improve interpretability and reliability and may reduce the demand for experimental training data by leveraging synthetic data generated from optical models.

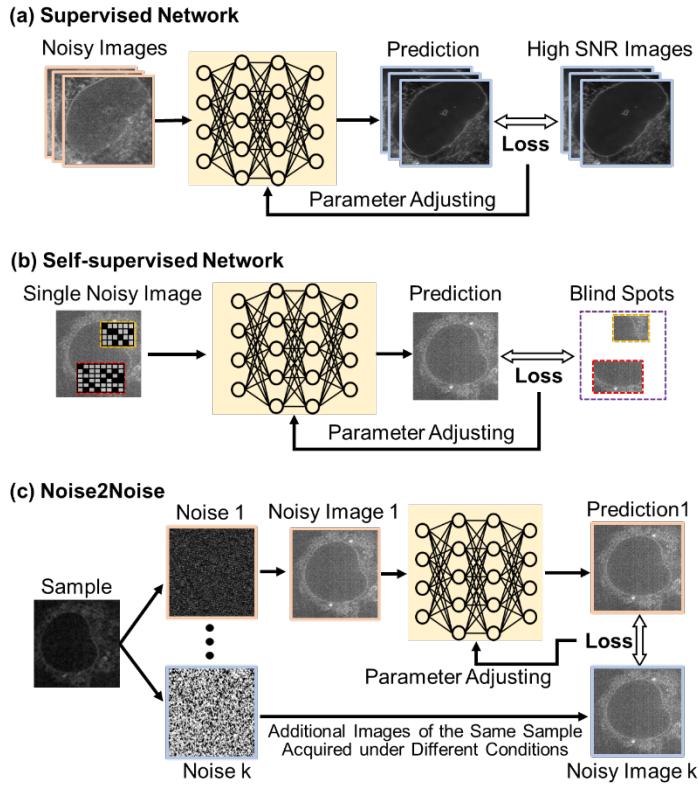


Figure 6.2: **Different denoising methods for microscopy images.** (a) Supervised denoising methods[117, 21], train a neural network using paired noisy and high-SNR images. (b) In self-supervised methods[54, 8], a portion of the image is masked (blind spots or black pixels in yellow and red subregions) for training, removing the need for explicit high-SNR images. (c) The Noise2Noise network Lehtinen et al. [58] is another form of supervised training, whereby the network is trained with a set of noisy images of the same underlying sample.

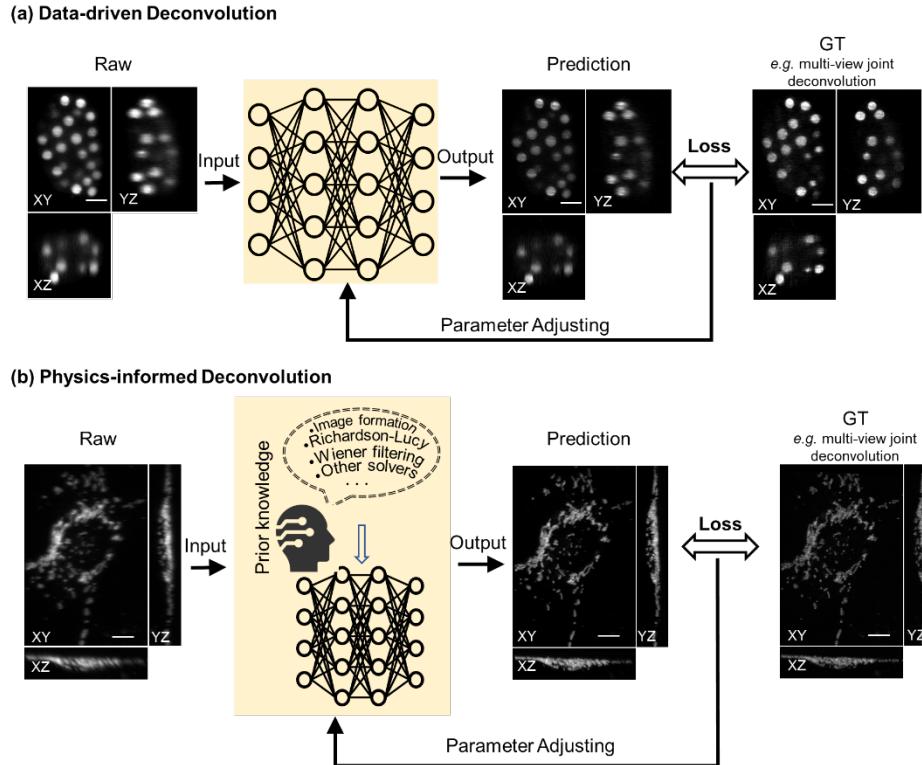


Figure 6.3: **Different Deep Learning-Based Methods for Deconvolution in Microscopy Images.** (a) Data-driven based deconvolution methods[35, 117], where a neural network is trained using only a large-set of paired raw data and ground truth (GT, e.g., joint deconvolution of multiview data). (b) Physics-informed deconvolution[60, 121], whereby a neural network is trained using not only paired datasets but also incorporates prior physical knowledge, such as embedding the image formation function in the loss function and/or the Richardson-Lucy framework or Wiener filtering in the network architecture.

### 6.2.3 Deaberration

Deaberration addresses image degradation caused by optical distortions due to refractive index mismatches in biological samples or imperfections in the optical system. These aberrations degrade image quality, particularly in thick or scattering samples. Recent advances in deep learning have introduced powerful alternatives to traditional AO for both explicit wavefront estimation and the prediction of cleaner images in which aberrations are suppressed (Figure 6.4).

Initial efforts focused on estimating the distorted wavefronts with AI, followed by explicit wavefront correction with hardware (e.g., a deformable mirror, or SLM) to obtain a clean image. The wavefront aberration can be decomposed as a sum of [Zernike polynomials](#). The problem of wavefront estimation then becomes inferring the amplitudes of different Zernike polynomials Saha et al. [92]. This approach combines the strengths of traditional AO with artificial intelligence [49, 48, 27]. Following the concept of deconvolution, researchers have also used deep learning to predict the aberrated wavefronts or blurring kernels (aberrated PSFs), also combining this information with post-processing deconvolution algorithms for image restoration[127, 83].

In contrast to explicit wavefront prediction, other efforts have trained neural networks directly on paired aberrated and corrected images [34, 43]. These models are designed to learn the mapping between aberrated and corrected images, enabling aberration correction as a purely computational post-processing step. This approach eliminates the need for hardware-based wavefront sensing or correction, making it a practical and scalable solution for many microscopy applications. More recently, a unified deep learning framework for simultaneous denoising and deaberration in fluorescence microscopy has also been developed Hou et al. [42]. By addressing both noise and aberrations in a single model, this approach simplifies the image restoration pipeline while delivering high-quality results, highlighting the potential of multitask models to streamline microscopy workflows.

### 6.2.4 Resolution Enhancement

The spatial resolution of optical microscopy is fundamentally limited by the diffraction of light, a barrier described by Abbe's law. Deep learning methods can learn to predict what a higher resolution image might look like, given lower resolution input data (Figure 6.5). Such methods use artificial neural networks to infer high-resolution (HR) details from low-resolution (LR) fluorescence images, which can offer a sharper image without the need for super-resolution microscopes. Cross-modality super-resolution (CMSR) refers to a class of techniques that use information from different imaging modalities to enhance resolution. Models such as CNNs and GANs are trained to learn the mapping between these modalities. For example, researchers can use super-resolution (e.g., STED) images to enhance diffraction-limited (e.g., confocal) images Wang et al. [114]; or use models trained on expansion microscopy data to predict super-resolution images from diffraction-limited inputs. Fourier-based approaches (e.g. DFCAN

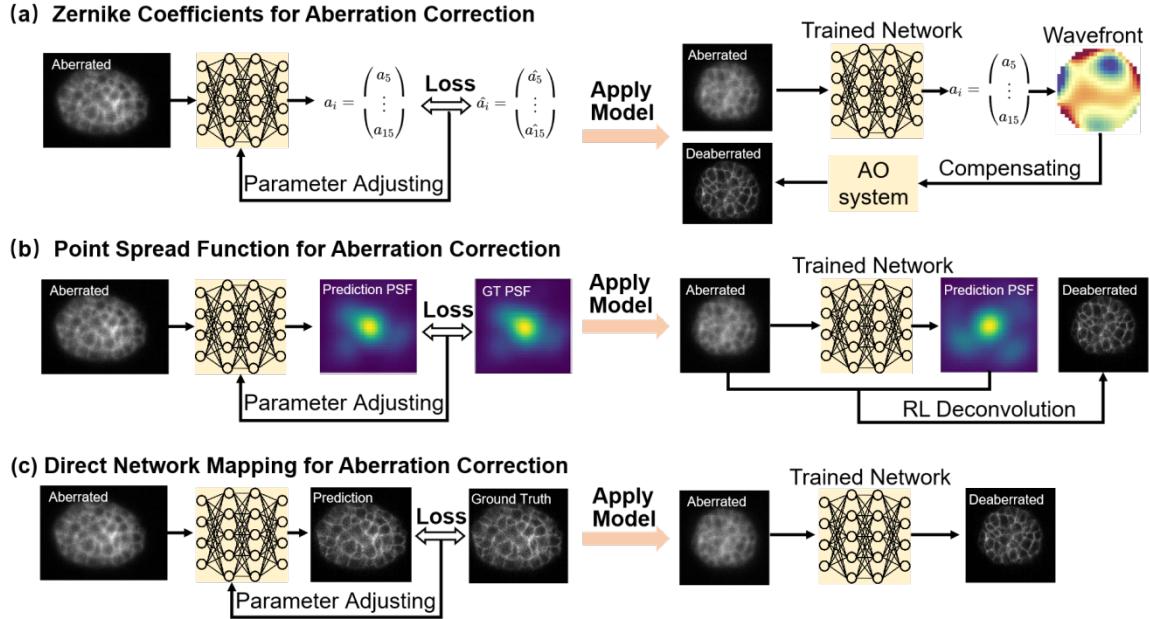


Figure 6.4: **Different Deep Learning-Based Methods for Aberration Correction in Microscopy Images.** (a) A method where the network is configured to learn the distorted wavefront (using a sparse representation of Zernike coefficients) [49, 48, 27]. The trained network is then used to predict wavefront distortions (in terms of Zernike coefficients) for a new dataset, followed by hardware-based correction using an adaptive optics (AO) system (e.g., a deformable mirror). (b) A method where the network is configured to learn the aberrated Point Spread Function (PSF) [127, 83]. The trained network is then used to predict the aberrated PSF corresponding to a new dataset, and a second step of image deconvolution is applied with the predicted PSF to remove aberrations from raw images. (c) A method where a neural network learns the mapping between aberrated and corrected images directly for post-processing aberration correction [34, 43]. The trained network is then used to directly predict the deaberrated image corresponding to a new dataset.

Qiao et al. [84]) leverage the frequency content difference across distinct features in the Fourier domain to enable the networks to learn the hierarchical representations of high-frequency information efficiently. While powerful, we note that these resolution enhancement methods offer a prediction at best and cannot truly retrieve higher resolution information absent in the low-resolution input data.

Another resolution enhancement task is the prediction of isotropic resolution from blurry axial views. The majority of fluorescence microscopy techniques, including most widefield, confocal, and light-sheet microscopes, suffer from unmatched resolution in the lateral and axial directions (i.e., resolution anisotropy), which severely deteriorates the quality, reconstruction, and analysis of 3D images. Deep learning models can be trained to predict and fill in the missing information in the z-direction using the higher lateral resolution view and knowledge of the anisotropic PSF, effectively making the resolution more uniform across the entire image[81, 75].

## 6.3 Practical Guidelines for Image Restoration

For researchers new to deep learning-based image restoration, successfully applying deep learning to image restoration requires careful preparation and a clear understanding of basic concepts. Here we briefly describe practical guidelines to help you get started with using deep learning for image restoration, from data preparation to model deployment. A step-by-step manual for different image restoration tasks (i.e., denoising, deconvolution, deaberration, and resolution enhancement) using RCAN is provided in Appendix A.

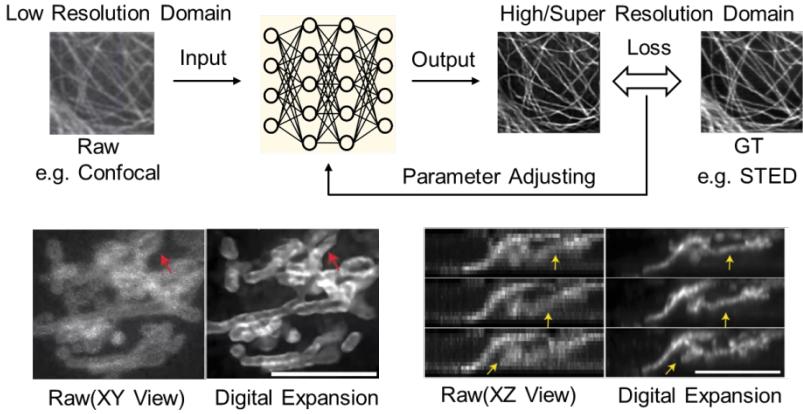
### 6.3.1 Considerations for Dataset Preparation

**Microscope Compatibility:** Ensure images are captured with consistent settings (e.g., exposure time, magnification). Common sources include confocal, widefield, light-sheet, or super-resolution microscopes.

**Paired vs. Unpaired Data:** Paired data (low-quality – high-quality pairs) is ideal for supervised learning. Acquire paired datasets of degraded and high-quality images for supervised training. In contrast, unpaired data is typically used for self-supervised methods (e.g., Noise2Noise). For unsupervised methods, collecting degraded images with varying conditions (e.g, different noise levels) helps the model generalize better across different conditions.

**Training Data Size:** According to the size of the acquired data, start with at least dozens of datasets of size  $128 \times 128$  for 2D and  $128 \times 128 \times 128$  for 3D. Augmentation with rotations, flips, and noise injections can expand your training dataset (see also Chapter 5).

**(a) Cross-modality Super Resolution**



**(b) Axial Resolution Enhancement**

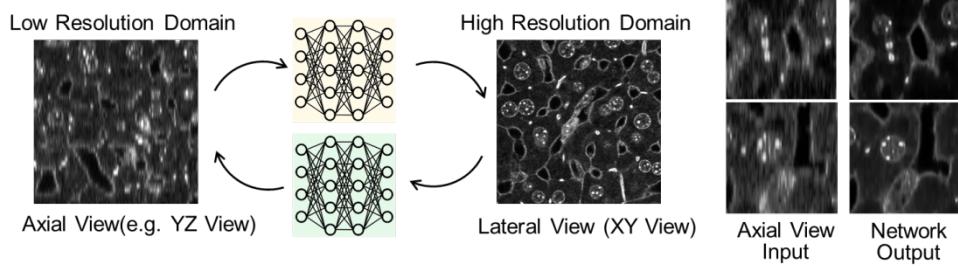


Figure 6.5: **Schematic of resolution enhancement with cross-domain networks.** (a) Cross-modality super resolution[21, 114]. Top row: the method takes diffraction-limited data (e.g., confocal microscopy data) as input and super-resolved data (e.g., STED microscopy data or expansion microscopy data) as ground truth. By training a neural network with paired datasets, the network parameters are updated based on the loss between the network’s output and the ground truth, thereby achieving a mapping from the low-resolution domain to the high-resolution domain. Bottom row: similar cross-modality learning can also be adapted using expansion microscopy as the ground truth. The red and yellow arrows highlight improvements in lateral and axial slices due to the restoration process. By ‘digital expansion’, we mean the super-resolution prediction based on expansion microscopy ground truth. (b) Axial resolution enhancement [81, 75]. The method uses datasets composed of low-resolution axial slices (sometimes synthesized from higher resolution lateral views Weigert et al. [117]) and high-resolution lateral slices, employing a one-step Park et al. [81] or two-step Ning et al. [75] framework to learn the mapping from the low-resolution domain to the high-resolution domain, thus achieving axial resolution enhancement for fluorescence microscopy 3D stack data.

### 6.3.2 Choosing a Network Architecture

The choice of network architecture can play an important role when performing image restoration (see also Chapter 4). Here are some popular architectures used for this purpose:

- Convolutional Neural Networks (CNNs): CNNs are currently commonly used because they are highly effective at capturing spatial features in images. For example, you can use U-Net Ronneberger, Fischer, and Brox [88], a popular CNN architecture designed for image segmentation that is also well-suited for restoration tasks.
- Generative Adversarial Networks (GANs): GANs can be used for tasks like image generation or synthesis (e.g., resolution enhancement), where the goal is to generate realistic images from degraded inputs. GANs consist of two networks: a generator (which creates images) and a discriminator (which evaluates the quality of the generated images).
- Autoencoders: These are unsupervised learning models that can be used for denoising, inpainting, and compression. Autoencoders consist of an encoder and a decoder that map input data into a lower-dimensional representation and then reconstruct it.
- Choose or design networks (e.g., CARE, Noise2Void), or even start with pre-trained models for faster deployment.

### 6.3.3 Training

Once you have your data, hardware, and software environment and have chosen a model architecture, it's time to train the model (see also Chapter 9). Training involves feeding the model the degraded images and their corresponding high-quality ground truth and then updating the model weights to minimize the difference between the restored and the ground truth images.

**Loss Functions:** Common loss functions for image restoration include mean squared error (MSE), structural similarity index (SSIM), and perceptual loss, which helps preserve perceptually important image details.

**Optimization:** Use optimizers like Adam or SGD (Stochastic Gradient Descent) to update the model's parameters during training.

Training deep learning models can take a significant amount of time, depending on the size of the dataset, the complexity of the model, and the hardware you use. You can use techniques like early stopping Ji, Li, and Telgarsky [47] to avoid overfitting and ensure the model is trained for an optimal number of epochs [94, 70].

### 6.3.4 Validation and Deployment

After training, it's important to evaluate your model's performance (see also Chapter 10). This can be done by testing it on a separate validation set (with ground truth reference) or test set (without ground truth) that wasn't used during training.

Common evaluation metrics for image restoration include peak signal-to-noise ratio (PSNR), SSIM, and Visual Quality Assessment. These metrics help assess how close the restored images are to the ground truth in terms of quality and structure. Perform biological validation to ensure restored images are artifact-free, such as expert evaluation, cross-validation with alternative techniques, or repeat experiments to evaluate uncertainty.

If the initial results are not satisfactory, you may need to fine-tune your model. This can involve: adjusting the model architecture (e.g., adding more layers or changing the number of filters); using a different loss function; and fine-tuning hyperparameters like the learning rate.

Once your model has been trained and evaluated, the next step is deploying it. Depending on the use case, you might deploy it for real-time image restoration or integrate it into a larger image processing pipeline.

## 6.4 Limitations and Future Perspectives

The integration of artificial intelligence (AI) into fluorescence microscopy image restoration has opened new avenues for biological discovery. However, as the field evolves, critical challenges and opportunities must be navigated carefully. Below, we outline key failure modes, caveats, and emerging trends, including the role of [transformers](#) and large models, to guide future research.

### 6.4.1 Caveats of AI-based image restoration

While AI has made remarkable strides in fluorescence image restoration, there are still failure modes that researchers must be aware of. One potential issue is overfitting, where a model becomes too specialized to the training data and struggles to generalize to unseen data. This can result in poor restoration performance on images that differ from the dataset used for training.

Additionally, loss of fine details [84, 102] and artifact generation [124, 64] remain a concern. Restored images may exhibit blurring, distorted textures, or unrealistic artifacts Bouchard et al. [14], especially in high-resolution regions or when attempting to restore challenging image regions (e.g., dense and overlapping structures Qiao et al. [85] or highly degraded structures Hou et al. [42]). For example, while doing confocal-to-STED microscopy restoration with RCAN, certain microtubules evident in the STED remain unresolved in the RCAN result, and RCAN

prediction for nuclear pores revealed slight differences in pore placement relative to STED ground truth Chen et al. [21]. While denoising fly wing data with CARE, the same networks predicted obviously dissimilar solutions across multiple trained models Weigert et al. [117]. During denoising, conflicts between restoring local details and enhancing global smoothing can arise [126, 76]. Due to AI’s inability to achieve a global understanding of semantic information, stitching artifacts may arise during large-scale data processing involving tiling operations Park et al. [80], resulting in inconsistent intensity, geometry, and textures. These artifacts may not always be immediately noticeable but can significantly affect subsequent analysis and interpretation. Ensuring algorithm robustness across varied datasets and imaging conditions is key to minimizing these issues.

Deep learning models often operate in a “black-box” manner. This means that while these models can produce useful predictions, the reasoning behind their predictions is not easily interpretable by humans. This lack of transparency and interpretability means researchers might not know exactly how or why certain features in an image are being restored in specific ways. This issue can undermine trust in the model’s results and limit its acceptance in critical applications.

AI-based image restoration techniques are highly dependent on the quality of training data. For fluorescence images, obtaining high-quality ground truth data can be challenging, especially when using highly specific imaging conditions or performing imaging on rare biological samples. The availability and variability of the resulting data can limit the model’s ability to generalize to diverse datasets.

Deep learning models can be computationally expensive during training and require powerful hardware for training and inference. This can be a significant barrier for some research labs or organizations with limited access to high-performance computing resources. Researchers need to consider these limitations when planning AI-based restoration projects and balance trade-offs between model complexity, data availability, and computational requirements.

#### 6.4.2 Outlook for the Future

With advances in transfer learning and self-supervised learning, AI models are likely to become more efficient, requiring less annotated data and computational power. As datasets grow and become more diverse, the dependency on highly specialized ground truth data will be reduced. Moreover, the integration of AI with microscopy platforms in real-time will enhance the ability to process images on the fly, providing immediate feedback to researchers and enabling dynamic imaging of live cells and tissues.

Newer methods of AI will be developed. Transformer models and large-scale models have shown remarkable success in other domains like natural language processing and computer vision. Transformer architectures, known for their ability to capture long-range dependencies and global context, could significantly improve image restoration tasks by better handling complex, large-scale image structures. Large models—trained on massive datasets—are likely

to offer even greater performance, able to generalize across different imaging modalities and restoration tasks. As computational resources continue to expand and more sophisticated models are developed, we can expect these methods to further push the boundaries of image restoration, achieving even finer levels of detail and more accurate reconstruction.

In conclusion, the future of AI in fluorescence microscopic image restoration is bright. While challenges such as interpretability, data quality, overfitting, and computational demands remain, the field is poised for rapid advances with new architectures and training techniques.

# 7 Adding AI to Your Hardware

An Introduction to Smart Microscopy

- Introduction
  - Define and motivate the “problem”
    - \* What is a Biological “Event”?
      - Why is it important to study these?
      - How is it distinct from an “Object”?
- Setting the bounds
  - What kinds of signals are we looking to extract events from?
  - What spatiotemporal scales are relevant?
  - What types of microscopes/imaging assays are in the scope of this chapter?
    - \* Limiting the scope
- Brief history/evolution of automated event/object detection in general
  - Offer some background to current methods
    - \* Including classical up to SVM
      - e.g micropilot and limitations
      - Case Study #1 : CellProfiler
      - Case Study #2 : Micropilot
- The need for new methods
  - Deep learning based and ML approaches
    - \* Where the algorithm learns what is important
    - \* What are the advantages of these methods?
      - What has been done recently in the literature?
    - \* What is required to label, train and implement these methods?
      - Special considerations for event detection modes!
      - Specifics of event detection in contrast the other ML tasks
- Real-time vs a posteriori inference

- Challenges, opportunities
- Event detection as the first step in the microscopy workflow
- Limitations and notes of caution related to inference
  - “Trusting the algorithms”
- Conclusion

# **Part III**

# **Image Analysis**

# **8 How do you select and find a tool?**

## **8.1 Introduction**

We are now transitioning away from a discussion of hardware and into a discussion of software. And, more specifically, in this chapter we focus on image analysis, particularly the human thought processes and decisions needed for selecting what software to use and figuring out how to appropriately/best use it. Amazing technological developments have occurred since the invention of microscopy almost half a millennium ago; Our current relationship to biological microscopy is profoundly shaped by the development of both quantitative microscopy and artificial intelligence in the mid-twentieth century. However, underlying the actual use of any technological development is a human (in our case, likely a biologist) selecting a specific technology to solve a specific problem.

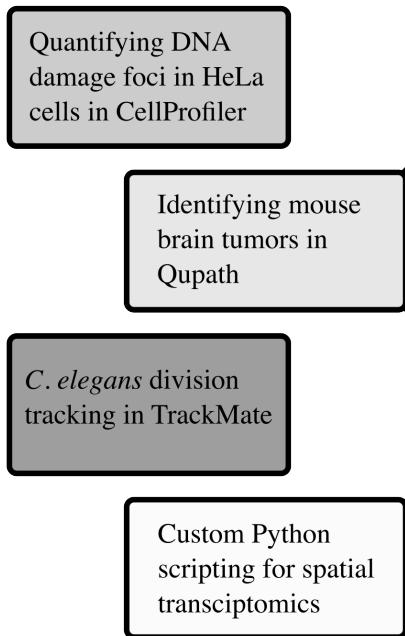
In this chapter, we answer the questions of how do you find new-to-you AI models and how do you assess whether a new-to-you model or tool will meet your needs? In this chapter, we first help you assess your needs. We then introduce a way of categorizing tools and help you use your needs assessment to select the right tool category. We suggest several places where you can find AI models for bioimage analysis and describe how to assess how well tools in those locations meet your needs. Finally, we show a couple of case studies that fulfill different requirements.

## **8.2 Assessing your requirements for a tool (in a perfect world)**

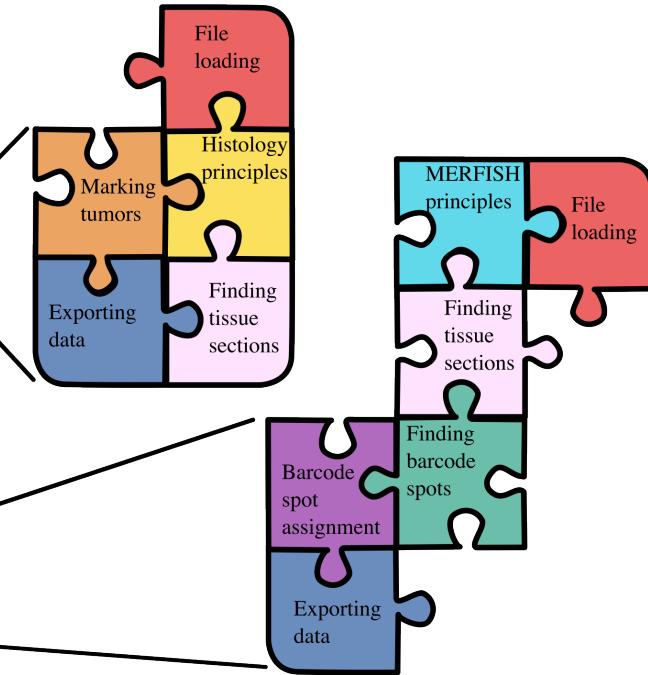
Before finding and selecting an AI tool, you first need to decide on the type of tool you need. Your tool selection will be influenced by a number of factors and you should first honestly answer the questions that follow.

Figure 1: Workflows contain modular components

**Image analysis workflows are not disparate monoliths...**



**...but are instead composed of many interchangeable and/or related components.**



**Figure 1.** When thinking of your workflow and how you might incorporate an AI tool, consider that a workflow (left of image) typically consists of multiple, modular subcomponents (right of image). You may be able to replace your whole workflow with an appropriate AI tool or you may be able to replace only a subcomponent while still keeping many of the original workflow subcomponents the same. Replacing a single subcomponent with an AI tool may force you to change surrounding subcomponents in your workflow - either with new tools or new configurations of existing tools - as the AI tool may require different inputs and produce different outputs.

Though you may have built a workflow across multiple softwares, it is worth considering whether a single software (or a reduced number of softwares) can run your workflow and part of this consideration includes assessing how well your desired AI tool can run from within other software. Some examples of softwares that can be used to run other softwares include CellProfiler Stirling et al. [106], Fiji Schindelin et al. [98], Bilayers Shah, Gogoberidze, and Cimini [101], ZeroCostDL4Mic Chamier et al. [16] or DL4MicEverywhere Hidalgo-Cenalmor et al. [41], and Jupyter notebooks Kluyver et al. [50].

Figure adapted from Haase et al. [36].

### **8.2.1 Questions about your overall workflow:**

- What is your desired output?**

You must be able to concisely and specifically state what you would like output by your workflow. Do you need to segment objects? If so, do you need [semantic](#), [instance](#), or [panoptic segmentation](#) (Box 1)? (see Chapter 4 and Chapter 9 for more information) Do you need measurements and if so should they be made on individual objects or on images?

- How will you assess your outputs?**

Will your assessment be qualitative or quantitative? Do you have ground truth? Do you have the expertise to assess output quality? (see [?@sec-10](#) for more information)

- What is your desired “quality” of your outputs?**

Are you expecting outputs that approach ground truth or will “quick and dirty” be enough for you?

- What does your ideal workflow look like?**

Do you need all of the outputs to be made in a single software or are you comfortable using multiple softwares? If your images require preprocessing, does that need to happen in the same step? The same software?

#### Ground Truth

**Ground truth** can mean many things. The Broad Bioimage Benchmark Collection (BBBC) Ljosa, Sokolnicki, and Carpenter [65] provides an nice example of this. In the BBBC, all datasets include ground truth, but that ground truth can be one or many of counts, foreground/background, outlines of objects, biological labels, location, or bounding boxes.

### **8.2.2 Questions about your resources: time, expertise, and compute:**

- How much time are you able to put into the task?**

An honest assessment at the beginning of any project about the time you are willing to invest is critical. If you are in a rush, you’ll probably want to select a method that is already familiar or best matches your existing competencies.

- What is your priority?**

Determining what your priority is goes hand in hand with assessing the amount of time you can put into your task. Perhaps you’re in a time crunch and speed is the most important consideration for you. Perhaps you have a specific [quality control \(QC\)](#) metric and you need the tool that will give you the outputs that optimize this metric. Perhaps you’ve always wanted to gain experience with a particular class of tool, so you want to

figure out if this is the right use case for trying it.

- **What is your level of computational domain expertise and comfort?**

If you don't have a high level of computational comfort, do you have the time and motivation to expand your skillset by building new skills outside of your comfort zone? Do you have a computational collaborator and how much time, in either teaching you or in handling the data themselves, are they able to contribute? Computation domain expertise has two critical, but separable, components. The first is the ability to understand what you are doing thoroughly enough that you can design, analyze, and interpret your experiment. The other component is the ability to comfortably interact with software as a computationalist might (e.g. are you comfortable working from the command line (which is text only) or would you prefer a GUI (graphical user interface), where you can point and click).

- **How do you like to learn things?**

Are you a person who likes to jump in and just turn all the knobs and buttons to see what they do? Do you love to cozy up with a good written manual at the end of the day? Do video walkthroughs bring you joy or frustration? How important is it to you that you be able to ask the tool creator questions directly? The type of documentation and instructional materials as well as their depth can vary greatly between tools.

- **What is your level of biological domain expertise?**

Are you confident that you fully understand the task so you can assess how well your AI model is performing? Do you have a biologist collaborator and how much time are they able to contribute to designing the experiment and/or analyzing data? Do you understand what controls you will need and/or corrections (such as single-color controls and/or measured flat-field corrections) you will need to make to be able to interpret your outputs?

- **What access do you have to compute resources?**

Do you need to be able to run everything on your laptop or do you have ready access to a high performance cluster or scalable cloud compute? Do you have access to GPUs? Different tools have different compute requirements, especially as you scale them, and those requirements don't always scale linearly with data size.

### **8.2.3 Questions about your input data:**

- **How difficult are your images?**

Perfect, clean data is the ideal input to any analysis. But that's not always the regime we're in. There are many sources of "difficult". A couple examples and corresponding questions are below.

- Do your images have debris or other technical artifacts such as out-of-focus regions?  
Do those artifacts need to be identified and removed? If identified, should the whole

image be dropped or do you need to keep non-artifact areas of the image?

- Do you have **metadata** that needs to be associated with your images? Is that metadata organized? How is that metadata organized (e.g. in the file name, in a .csv file, in a picture of a handwritten lab notebook) and does it play nicely with the tool you would like to use?
- Are your images in a friendly format? Are they in a proprietary file format? Are they in a file format that allows access to chunks?

- **How big is your data?**

The larger data is, the harder it can be to work with. If your data is large, what dimensionality is big? e.g. many images, large individual file sizes, many channels

- **Do you need outputs in multiple dimensions?**

If you have z-planes, do you need 3D objects or can each plane be handled separately? If you have multiple timepoints, do you need objects tracked across timepoints or can each timepoint be handled separately?

- **Do your images require preprocessing?**

There are many different reasons that images might require preprocessing and some of those reasons may be a way to overcome concerns/technical challenges raised in other questions above. Some examples of preprocessing include stitching or cropping of images, denoising, background subtraction, or flat field correction.

#### Box 1: Segmentation methods

In computer vision, there are several discrete kinds of segmentation, although in the field of bioimaging we often refer to all kinds of segmentation under the single blanket term of “segmentation”. **Semantic segmentation** divides an image into classes. Ilastik is an example of a popular image analysis software that performs semantic segmentation. **Instance segmentation** detects individual, specific objects within an image. CellProfiler is an example of a popular image analysis software that is most commonly used for instance segmentation. **Panoptic segmentation** is a combination of semantic segmentation and instance segmentation that separates an image into regions while also detecting individual object instances within those regions. Deep learning can be used for semantic, instance, or panoptic segmentation. Most classic image analysis methods are built around instance segmentation for cell-based images, though semantic segmentation is not uncommon for tissue/histology images.

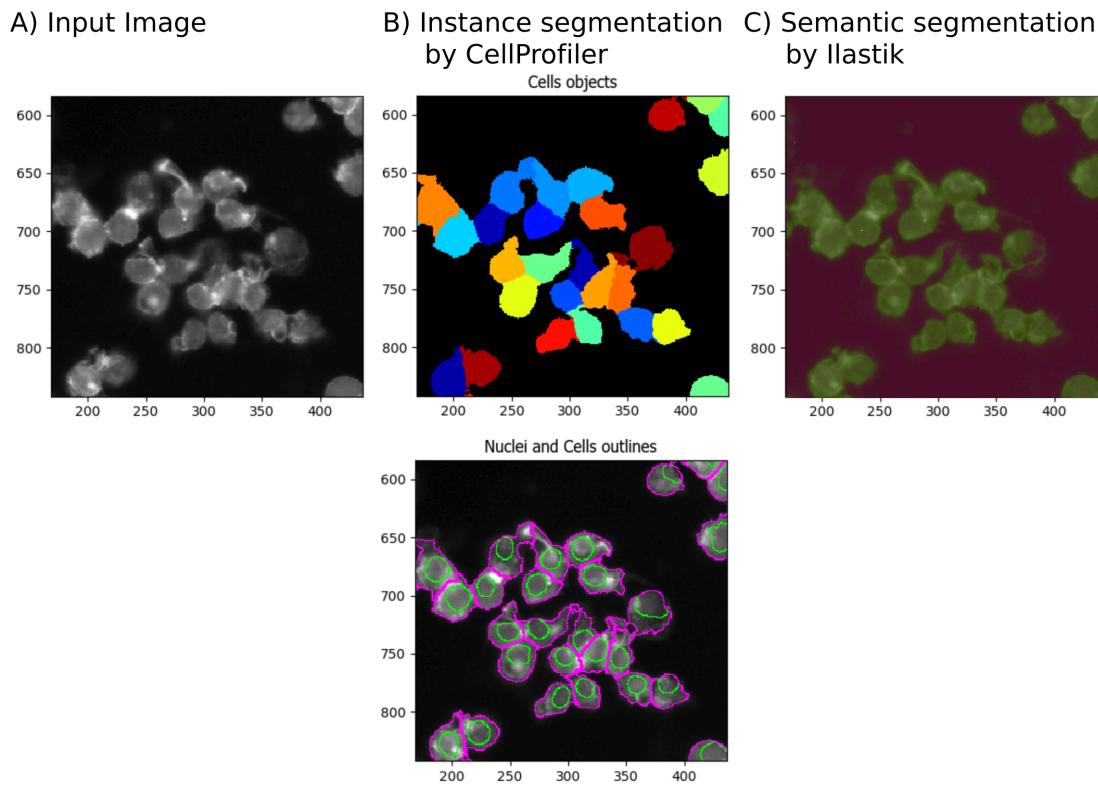


Figure 8.1: Segmentation comparison of CellProfiler and Ilastik

Example of different types of segmentations produced between CellProfiler and Ilastik. A) Input image of Drosophila Kc167 cells, provided in the Example pipeline packaged with CellProfiler. B) CellProfiler uses classic image processing to create instance segmentations. The individual objects identified are shown as individually colored masks in the upper panel. The lower image is the input image overlaid with nuclei objects outlined in green and cell objects outlined in pink.

C) Ilastik uses pixel-based machine learning to create semantic segmentation. The image is the input image with green shaded areas in the class “Cells” and red shaded areas in the class “Background”.

### 8.3 Assessing your requirements for a tool (in the real world)

In several previous surveys we have run, scientists have reported that they generally don’t actually follow any thoughtful, rigorous process for deciding what tool they will use for a particular process. Instead, they sit down with a thing they know how to use and hope they

can make it follow their use case Jamali et al. [45] Sivagurunathan et al. [105]. If they don't already have familiarity with a tool that is appropriate for their task, they ask their labmates or maybe do a quick internet search. So know that if you found the section above overwhelming, you are not alone.

If you already have familiarity with a tool and it fits your newest use case (especially if your newest use case isn't *very* different from a use case where you've previously successfully used it) go ahead and stick with that tool. "If it ain't broke, don't fix it" is a common adage for a reason. But even for existing use cases, it's worth actively contemplating your priorities every-so-often - maybe your GPU cluster is filling up these days and is harder to access, or after a new operating system upgrade a tool that was working great is now randomly crashing, or maybe there simply is something new (or new-to-you) available that will make your life easier. In our experience, the best way to be sure you're best serving your scientific needs is simply to step back and consider your big picture goals and whether your tool choices are serving those goals with some regularity.

Once you have conducted a thorough needs assessment, it should *theoretically* be simple to determine if a given tool or model is the right one for you. Unfortunately, no tool is likely to perfectly fit every single need that you have! You will likely need to break your needs into "must haves" and "nice-to-haves", and rank tools accordingly. Make sure you consider not just how well the tool will perform your specific task, but also how well it will fit in your workflow (Figure 1)

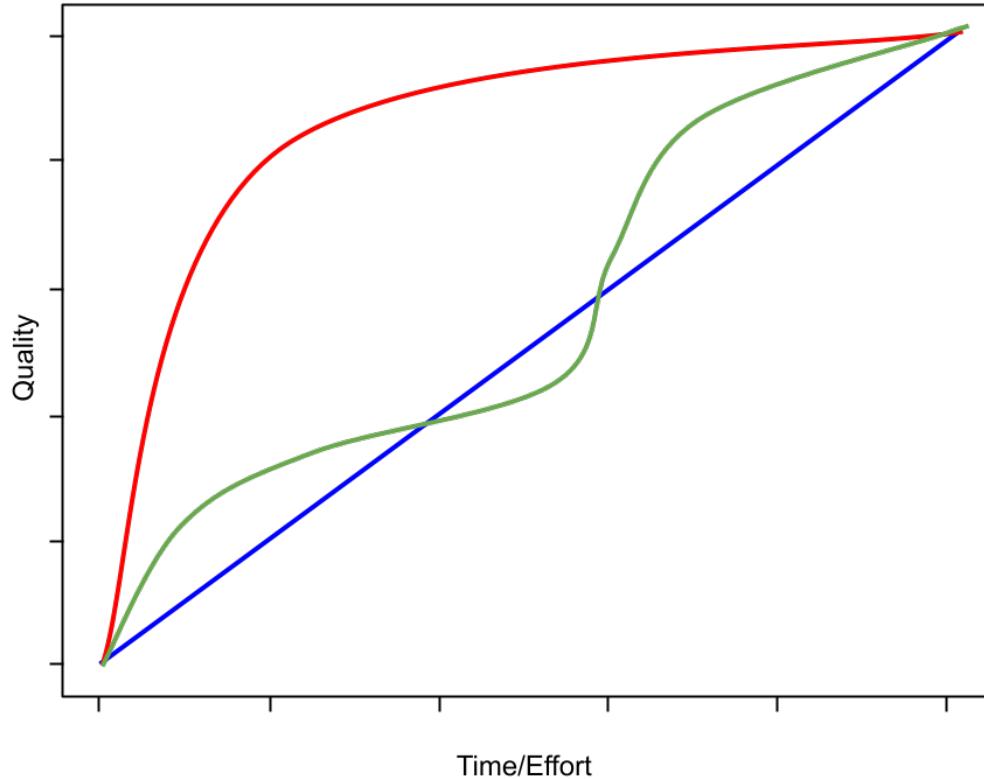
## 8.4 How do I decide when my workflow is “good enough”?

Deciding when a workflow is “good enough” can be really hard! We generally factor in two major considerations when deciding what quality level we are aiming to achieve: Is this a pilot or final data? and How big is the [effect size](#) we’re looking for?

We generally hold pilot data to a lower quality standard than final or production data. We are looking for a proof of concept that our assay/method can work but understand that there are likely to be refinements and improvements at many points in the workflow that will increase our final quality. Additionally, during the piloting phase, there are likely to be changes introduced that change the inputs to our workflow so time spent beyond a certain point of workflow refinement is lost with the retuning/adjustments required for the next pilot batch.

Phenotypic effect size has a large impact on necessary quality level. Simple, easily distinguished phenotypes (e.g. GFP is either in the nucleus or cytoplasm) may reach statistical significance with many fewer individual measurements and using much simple/cruder regions of interest (e.g. cell bounding boxes instead of careful cell margin segmentations) than more subtle phenotypes.

Research progress is rarely linear



**Figure 2.** As scientists who care about the quality of our work, it's tempting to always try to maximize the accuracy of our analysis. But it's worth considering that "accuracy" vs "time spent" is often an asymptotic curve - you can keep pushing the accuracy higher and higher, but the gains may be marginal past a particular point. Time and effort spent on a project can have very different relationships with the output quality. Though one might hope to be in a regime where the amount of time spent is linearly correlated with the quality of a project's output (blue line) so that it is clear exactly how much time must be put in to get a desired product, research rarely proceeds in such a fashion. Instead, it progresses in unpredictable fits and starts (green line) and/or it is in a state of diminishing returns (red line) where the closer you get to "perfect", the longer it takes to improve the quality.

In Cimini [22] we detail and provide examples for a few key heuristics that we think about when deciding if a workflow is "good enough". Those heuristics are:

1. How close to accurate do I need to be to assess the % or fold change I expect to see in this experiment?

2. How close to 100% accurate is it possible to get with my current workflow?
3. How important is this aspect of my experiment to my overall hypothesis?
4. How wrong is my current pipeline output?
5. What else could I do (e.g. develop a different approach/workflow, learn a new tool, generate new input data) in the time it will take me to make my pipeline maximally accurate?

## 8.5 Choosing the kind of tool to use

Let's consider the use case of instance segmentation (see Box 1), which is often a critical step in image analysis workflows. We have defined 6 categories of segmentation methods, each of which has its valid use cases. Ordered by increasing levels of computation comfort required they are:

- [Manual annotation](#)
- Classical image processing
- Pixel based machine learning
- Pretrained deep learning models
- Finetuned deep learning models
- From-scratch deep learning models

We will not cover from-scratch deep learning models in this chapter. They are covered in Chapter [9](#).

It is worth noting that many of the tool suggestions below are not confined to a single “class”. Many tools listed in the non-deep learning categories allow you to run pre-trained deep learning models inside them (e.g. CellProfiler has a RunCellPose module, ImageJ and QuPath also contain ability to run [pixel classifiers](#)). All deep learning models *can* be re-tuned with some effort (they may just not have a friendly interface for doing so).

Method category	Suggested light microscopy tools	Time required	Best if your priorities include	Biological level required	Ability to handle noise and variability	Largest drawback
<b>Manual Annotation:</b> Hand-drawing in boundaries of all objects	ImageJ Schindelin et al. [98] Schindelin et al. [99], Napari <i>napari: a multi-dimensional image viewer for Python</i> [74], see also <a href="#">this forum thread</a> for 3D annotation	Highest	High accuracy/ outputs approaching ground truth	High	High	Inter- and even intra-operator manual annotations (e.g. after time or image rotation) variability. Irreproducibility. Can not be automated.

Method category	Suggested light microscopy tools	Time required	Best if your priorities include	Biological level required	Ability to handle noise and variability	Largest drawback
<b>Classical Image Processing:</b> Combining mathematically well-defined operations such as thresholding and watershedding to create objects	CellProfiler Stirling et al. [106], QuPath Bankhead et al. [7], Icy Chaumont et al. [17]	Medium	Needing to have a good sense of what the failure cases will be (e.g. when cells are too big, too small, too crowded, etc); needing to run in minimal computational footprints	High	Initially low, can be improved with experience	Most performant on very bright objects on dark backgrounds

Method category	Suggested light microscopy tools	Time required	Best if your priorities include	Biological level required	Ability to handle noise and variability	Largest drawback
<b>Pixel based machine learning:</b> Using properties of individual pixels to classify each pixel into belonging to a given set of “classes” (e.g. “true signal”, “autofluorescence”, and “background”)	Ilastik Berg et al. [12], Labkit Arzt et al. Using [5]	Low to Medium	Identifying well-separated objects that are not bright-on-dark or are not the <i>only</i> bright-on-dark objects present	High	High if properly represented in the training set	Must be followed by Classical Image Processing if individual object masks are required

Method category	Suggested light microscopy tools	Time required	Best if your priorities include	Biological level required	Ability to handle noise and variability	Largest drawback
<b>Pre-trained deep learning models:</b> Using a set of learned transformations (created by the developer) to define the positions and boundaries of objects	Stardist Schmidt et al. [100], Instanseg Goldsborough et al. [30]	Low	Maximum likelihood for good quality with minimal human time required	High	High if properly represented in the training set (which you may not know)	Unlikely to work if data similar to your data is not present in the training set. Can be hard to package into workflows.
<b>Fine tuned deep learning models:</b> Updating or adding to a pretrained deep learning model with your data	Cellpose Pachitariu and Stringer [79], $\mu$ SAM Archit et al. [3]	Medium	Maximum likelihood for very good quality with less human time than manual annotation	High	High if properly represented in the training set	Most computationally intensive. Many available models are not provided in a non-computationalist friendly format. Can be hard to package into workflows.

You may have noticed that all methods are listed as requiring a high level of biological expertise. It can be tempting for researchers without deep biological expertise to assume that they could just use a pretrained or fine tuned deep learning model and call it a day. However, we would caution that, while they may require less of a biologist's time to tune at a first pass, it is still essential to involve a biologist with subject matter expertise to verify that the models are producing results that match known biology.

## 8.6 Finding tools and models

### 8.6.1 Where to look for tools and models

There are many places that one can look to find new (or new-to-them) tools and models. A few suggestions are as follows:

- BioImage Model Zoo Ouyang et al. [77] is a community-driven AI model repository that offers a variety of pretrained AI models. All models are described by a common Model Resource Description File Specification which simplifies using models in the BioImage Model Zoo with partner software such as ilastik Berg et al. [12], ImJoy Ouyang et al. [78], Fiji Schindelin et al. [98], deepImageJ Gómez-de-Mariscal et al. [31] and ZeroCostDL4Mic Chamier et al. [16].
- Bioimage Informatics Index (Biii) Zhang et al. [122] is a community-curated index that includes software and helps researchers find tools whether their search is problem-based (e.g. “find nuclei in cells”), method-based (e.g. “active contour-based segmentation”), or tool-based (e.g. “CellProfiler”).
- The Scientific Community Image Forum ([forum.image.sc](http://forum.image.sc)) Rueden et al. [89] is a platform where users can pose bioimage software-related questions and receive feedback from a broad user community and maintainers from >75 partner softwares.
- Academic literature can include biology-focused research where users describe applications or computer science research where the tools and models themselves are more likely to be described. Keeping an eye on preprint servers like arXiv (particularly in Computing Research Repository or Quantitative Biology sections) and bioRxiv can help you find the most cutting-edge developments.

#### Bioimage analysts

There is a whole discipline of people who specialize in figuring out workflows for analyzing microscopy images - their official titles often vary, but many think of themselves as *bioimage analysts*. If you want to find a local bioimage analyst, or even become one, you can check out the [Global Bioimage Analysts' Society \(GloBIAS\)](#), which holds events,

coordinates trainings (including free help sessions), and hosts a database of bioimage analysts and bioimage analysis trainers.

## 8.7 Case Studies

### 8.7.1 Case Study 1:

#### 8.7.1.1 Introduction

Elena is an image analyst who uses high content microscopy as a method for exploring drug mechanism of action. She uses an automated microscope to acquire 9 fields of view in each well of a 384-well plate and just acquired a batch of 4 plates worth of images. She often performs the same assay so she has a series of pipelines that she has nicely tuned for her usual experimental parameters. Elena has a new collaborator that would like to perform a similar screen in a different cell line as the collaborator is interested in a different specific area of biology. Because this is the first time that Elena has performed this assay in a new cell line, she needs to carefully tune her existing pipelines so that she has high-quality segmentation of Nuclei and Cell objects.

#### 8.7.1.2 Workflow

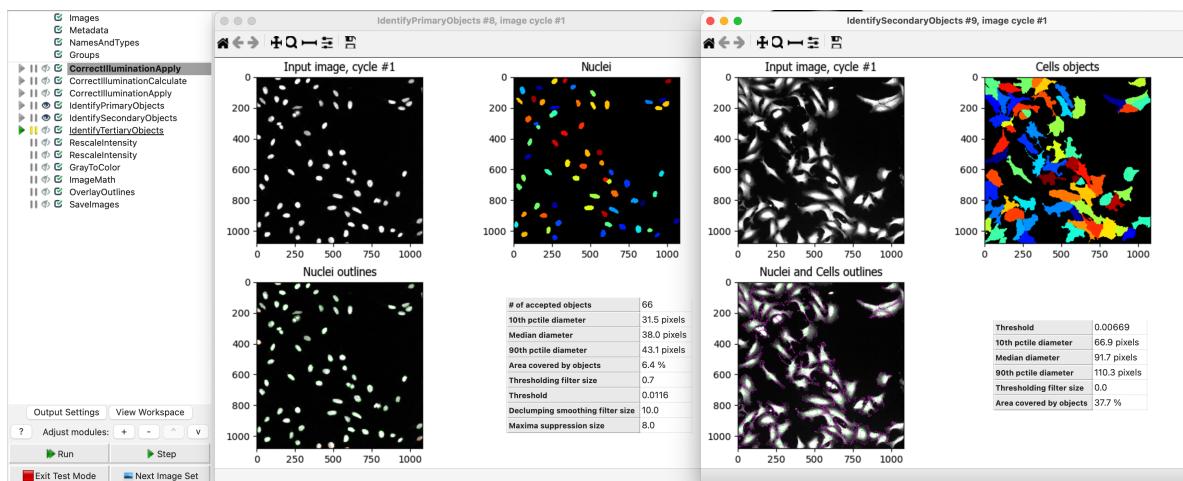


Figure 8.2: CellProfiler for nuclei and cell segmentation

Elena usually uses CellProfiler as an image analysis workflow tool, using classical image processing methods for object identification in her pipeline. Elena typically performs her

screens in A549, U2OS, or HeLa cells which are all considered easy to use for high-content microscopy because they are relatively consistent in size, relatively round, and can grow both sparsely and confluently while remaining in a monolayer. She's quite used to making parameter changes to the IdentifyPrimaryObjects and IdentifySecondaryObjects modules to fine-tune her nuclei and cell segmentations, respectively. The image above shows shows the modules in her CellProfiler pipeline (left) and the outputs of her IdentifyPrimaryObjects (center) and IdentifySecondaryObjects modules (right).

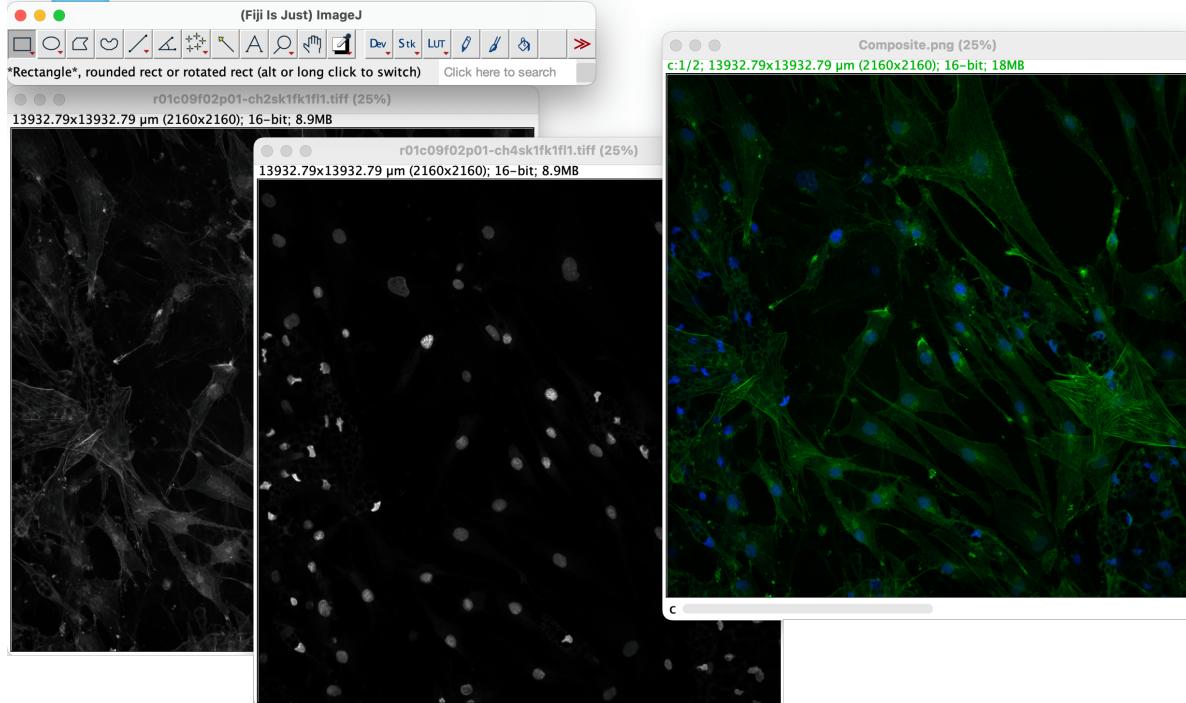


Figure 8.3: Fiji for compositing a new png image

Elena's collaborator has acquired images in an adipocyte cell line and they are morphologically quite different from her typical screening cell lines. After being unsatisfied with what she was able to tune using classical segmentation in her standard Cellprofiler pipeline, Elena decides to turn to an AI tool. She starts by opening Cellpose in its web console as she has heard from colleagues that it is both very good out of the box and allows a user to fine tune it for their own data. She initially struggles to load one of her own images into the web tool but notices that the web page says “upload one PNG or JPG <10 MB”. She also notices that the example images that are most similar to hers are in color, with a nuclear channel in blue and a cell marker in green. Elena's raw images are in .tif format with a single grayscale channel per image so she uses Fiji to make a composite .png image. The image above shows her single-channel .tif images opened with FIJI (left, center) and the composite .png image that she created (right).

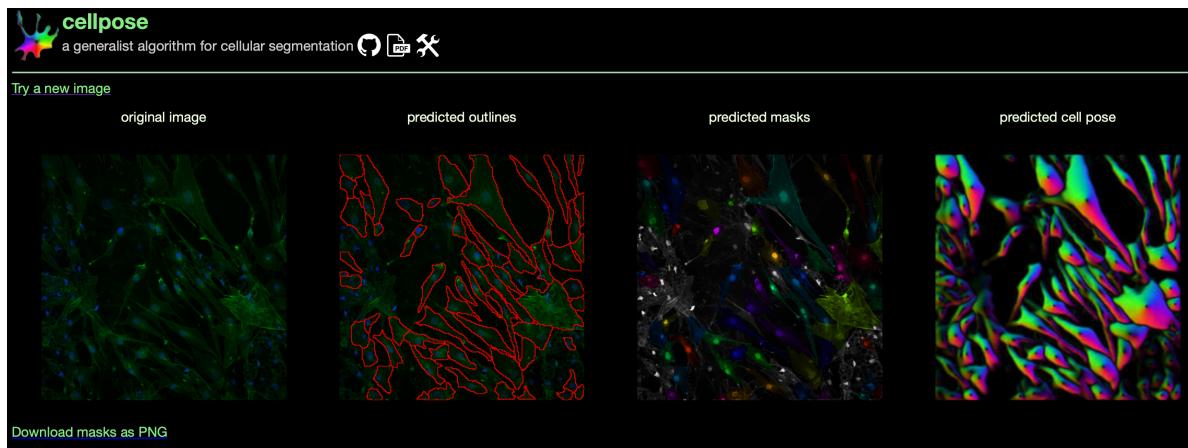


Figure 8.4: Cellpose web console

Elena loads her new composite .png into the Cellpose web console and sees that it has identified a number of reasonable cell objects but is not of the quality that she wants for use in her screen. However, Elena is excited to explore a new tool and has been actively building her computational skills so she decides to proceed with Cellpose and see if she can improve the outputs. The image above shows her example image loaded into the Cellpose web console (far left) along with the predicted object outlines (center left), masks (center right), and flows (far right).

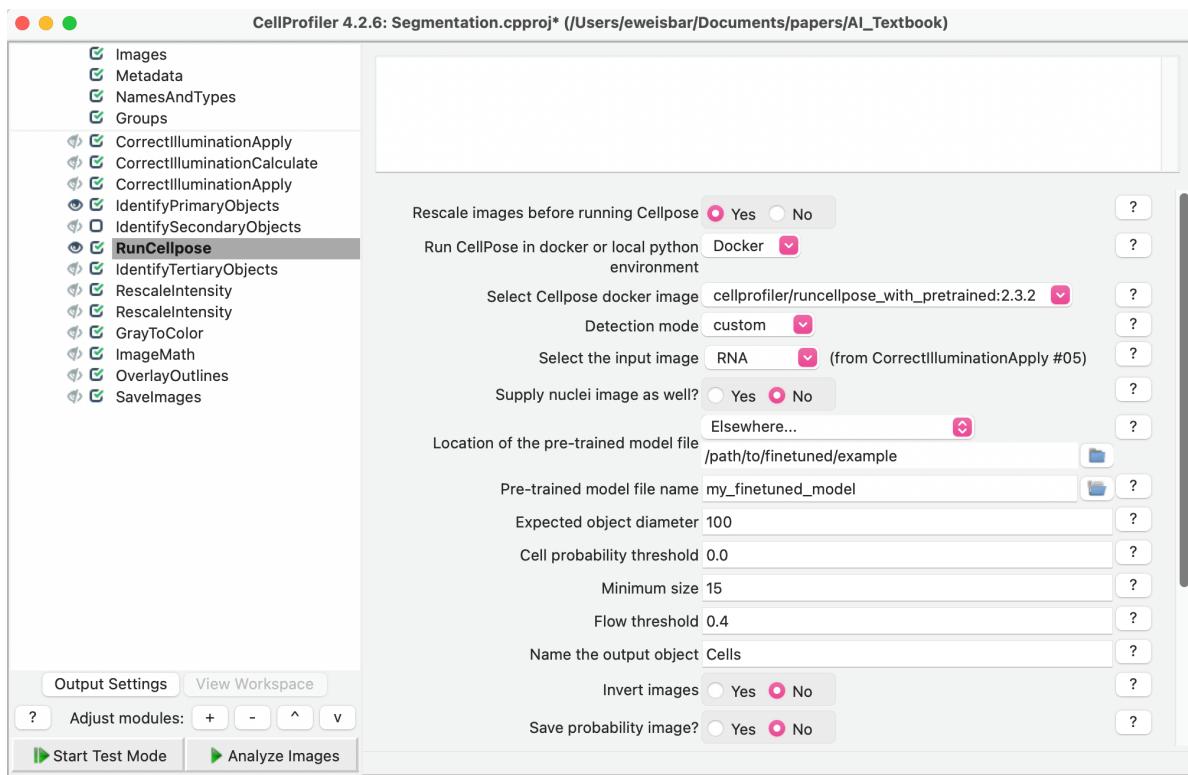


Figure 8.5: RunCellpose in CellProfiler

Elena takes advantage of the fact that Cellpose allows for fine-tuning of the pretrained deep learning models it provides. She re-trains Cellpose using a Colab notebook and sees that it now performs quite well on test images from her dataset. Because CellProfiler is also an image analysis workflow tool, Elena is able to insert a RunCellpose module in her original pipeline instead of the original secondary object identification module and keep the rest of her image handling and analysis the same while using the newer, better performing object identification that Cellpose provides. The image above shows her new CellProfiler pipeline with the RunCellpose module pointing to her custom fine-tuned Cellpose model.

#### 8.7.1.3 Conclusion

This case study starts with a scientist trying to use her standard classical image processing methods for object identification but finding that she gets better results with a new challenging segmentation task by moving to an AI tool. It shows iterative improvements using the same tool in different ways and highlights some of the simple changes she needed to make to her overall workflow to use the new tool.

## **8.7.2 Case Study 2:**

### **8.7.2.1 Introduction**

Esteban is an image analyst, working with imaging data acquired by a collaborator. The collaborator studies zebrafish brains and sent Esteban fluorescent images taken in a large z-stack in live zebrafish. The zebrafish express a GFP-tagged protein in their microglia and the collaborators would like to be able to identify the microglia using an automated pipeline so that the researchers can ask biological questions such as where they are within the brain and whether they are close to other cell types (labeled in other fluorescent channels). The task is complicated because there is additional autofluorescence in the skin that is the same intensity as the microglia and they do not want skin autofluorescence identified. The images also have technical challenges because of the nature of their acquisition: because they are 3D live images, as you go deeper in the tissue the signal gets dimmer. This makes using normal [computer vision](#) segmentation quite hard.

### 8.7.2.2 Workflow

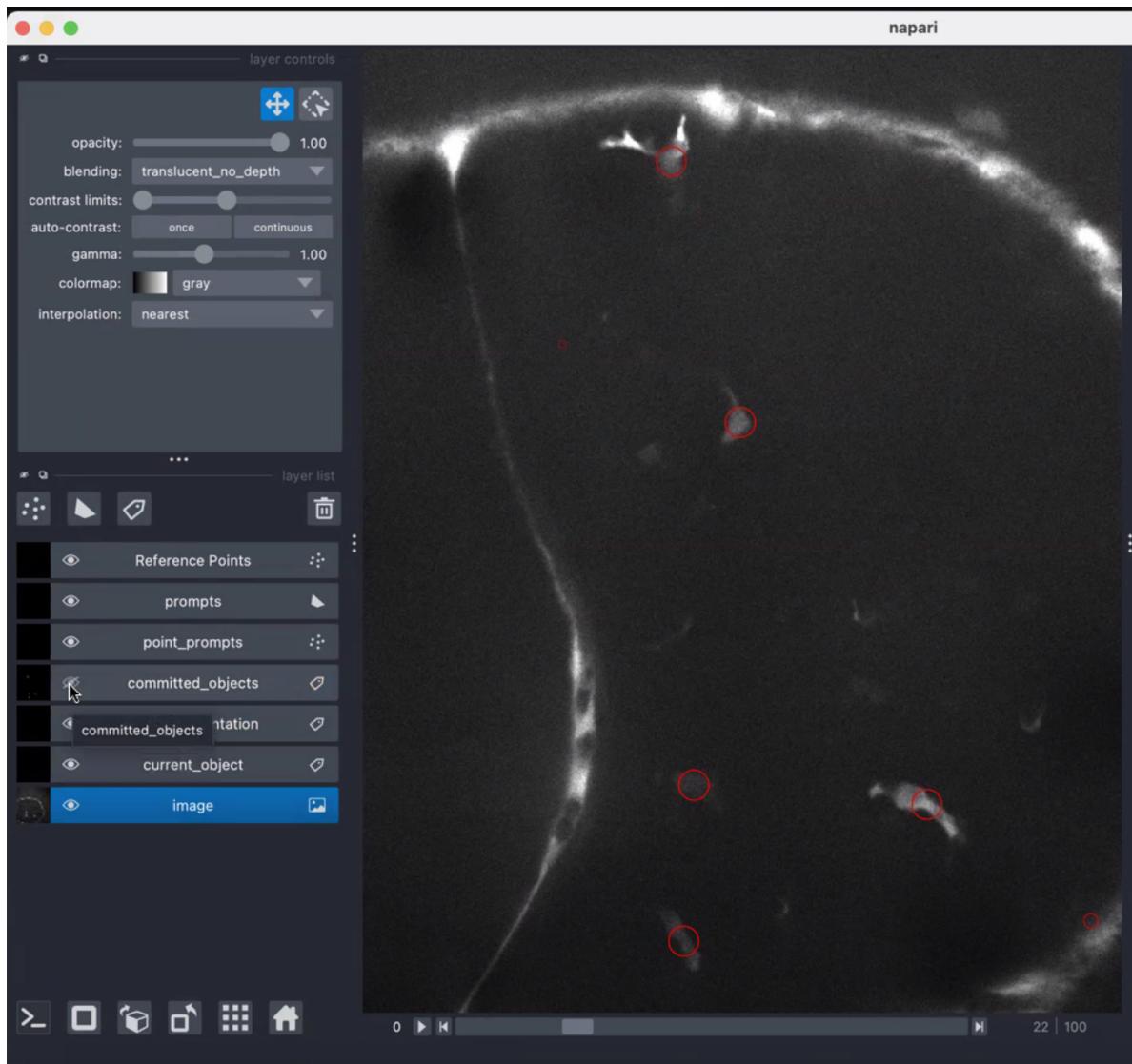


Figure 8.6: Point Annotations in napari

To help orient Esteban to the image data, he was given a spreadsheet of x,y,z coordinates of microglia that the collaborators manually annotated. Esteban started by opening the images and their point annotations in napari to orient himself to the data. The image above shows an example image opened in napari with red circles around the point annotations.

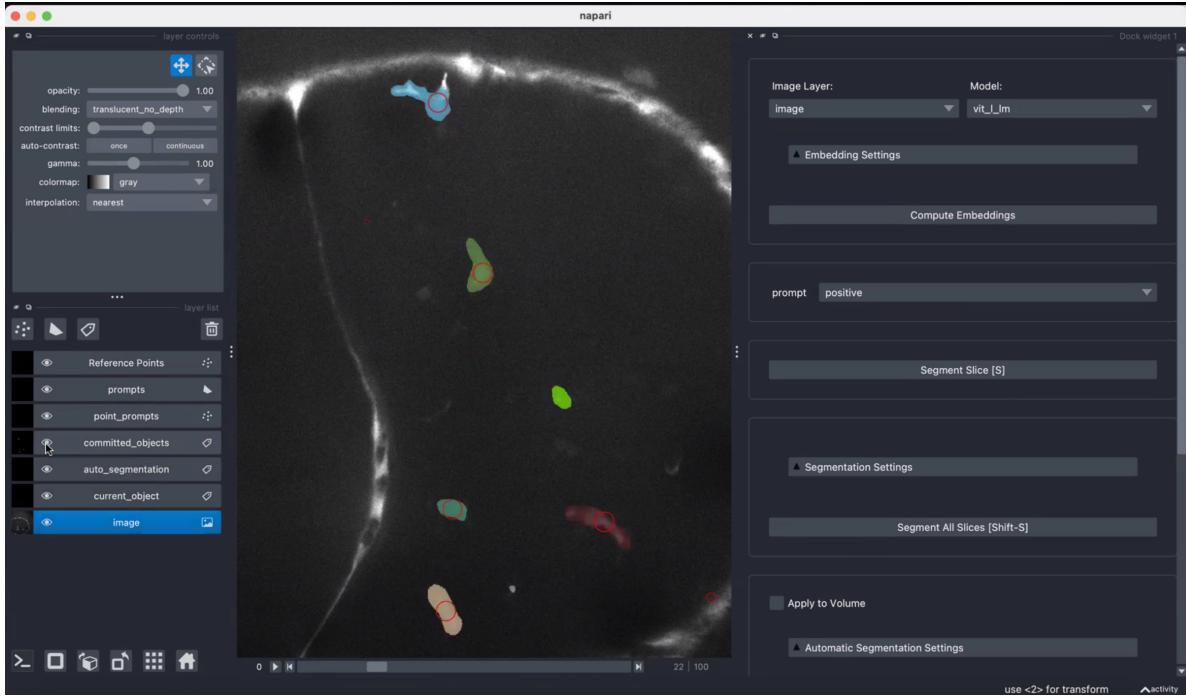


Figure 8.7: Guided Annotations in napari with SAM

Esteban then loaded  $\mu$ SAM in napari to use as an annotation tool so that he could use the point annotations to make object masks. He tried using automatic segmentation but it didn't work on his images. Instead, he used  $\mu$ SAM assisted segmentation, manually providing positive prompts, and got out reasonable segmentations of the microglia in the image. The image above is quite similar to the previous image but also has the  $\mu$ SAM segmentations in color overlays.

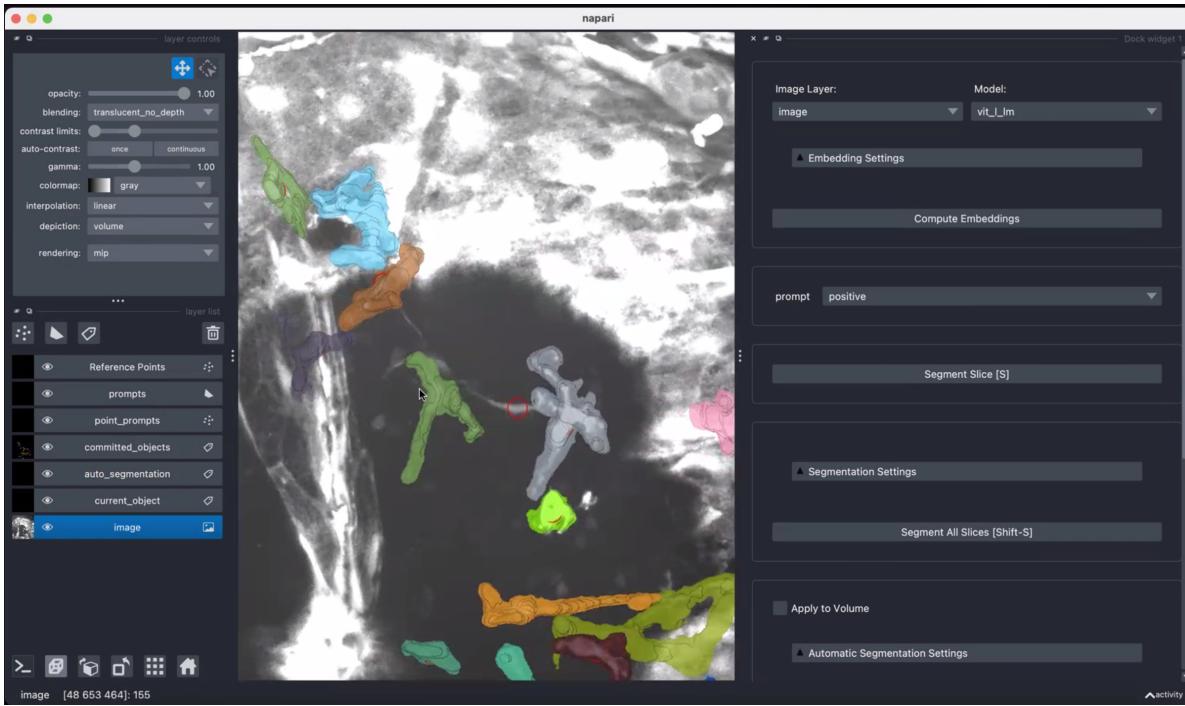


Figure 8.8: Visualizing 3D annotations in napari

Though Esteban has three dimensional data, the segmentations he is creating are 2D so he confirms that he is happy with his final masks by examining the image and masks in a 3D viewer. The above image shows a 3 dimensional view of his data that contains all of the image slices and annotations viewed together.

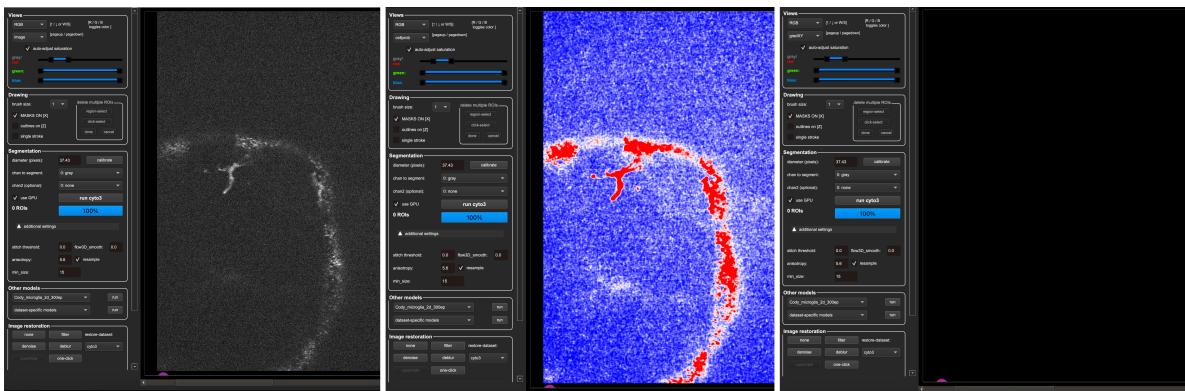


Figure 8.9: Cellpose predictions

Now that Esteban has both images and masks, he calls Cellpose in Python and retrains the cyto

model with his annotations. He runs the retrained model on a new image (left image) and finds that the cell probability prediction seems to identify microglia and autofluorescence (center image), but, unfortunately, the gradient prediction is not finding any cell objects (right image). Esteban's training images are quite sparse (the vast majority of the image is background) so the Cellpose model learned that it can be mostly correct by deciding that *everything* is background.

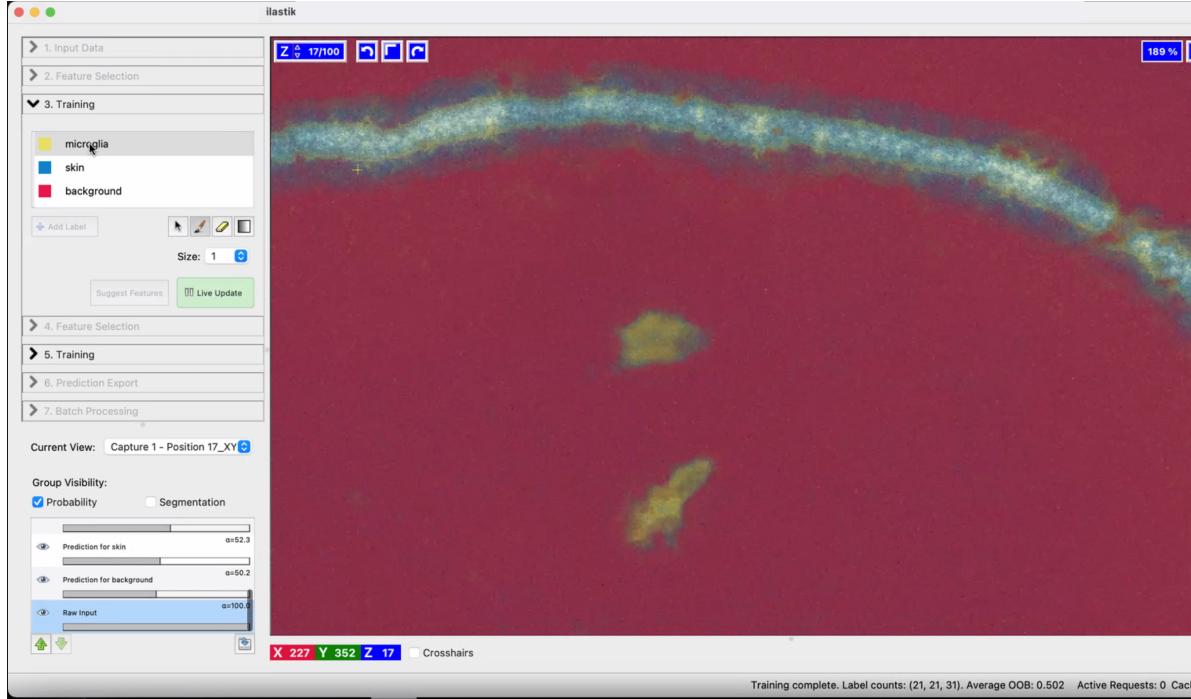


Figure 8.10: ilastik 3 class predictions

Since neither a pre-trained model ( $\mu$ SAM) nor a fine-tuned model (Cellpose) worked on his images, Esteban decided to try pixel classification. He used ilastik, first providing annotations for microglia and background. The classifier worked quite well at distinguishing microglia from background, but classified the skin autofluorescence as microglia. Esteban built on this by using larger filters so that the classifier had more contextual information, training 3 classes (microglia, skin, and background) instead of 2, and adding in autocontext Kreshuk and Zhang [52] which improves classification by running it in multiple stages. Esteban's final 3-class pixel classifier performs well to identify microglia separately from skin across the images in his datasets. The above image shows a zoom-in of one of his images with the three classes overlaid - red for background pixels, blue for skin pixels, and yellow for microglia pixels.

### 8.7.2.3 Conclusion

This case study starts with a scientist trying multiple state-of-the-art deep learning segmentation tools for a specific, challenging segmentation task. He ultimately discovers that a machine learning pixel classifier works better for this task, illustrating that 1) “less-advanced” tools can perform better than deep learning tools and 2) biological expertise is always needed when inspecting any tool’s outputs.

## 8.8 Conclusion

As we shift our focus from hardware to software, it is common and understandable to feel overwhelmed - the bioimage analysis tool space is a constantly-changing array of best practices, and not all microscopists are highly computationally comfortable. This chapter therefore mostly talks about *how to make decisions*, as opposed to *use X tool for Y task*. Ultimately, the “right tool for your task” is the tool (or chain of tools) that allows you to be confident that your analysis has been performed **well enough**, and helps you understand the answer to your biological question and how confident you can be in your conclusion. Even if you try a tool and decide it isn’t the right solution for this question, this time still isn’t wasted - it simply enlarged the toolbox you can use for answering future questions. In the subsequent chapters, we will teach you about how to train models which could be important parts of your toolbox, but we suggest you keep our guidelines in mind as you finalize the decisions around when your model is “trained enough” and how you decide on your overall workflow.

# 9 How to Train and Use Deep Learning Models in Microscopy

Practical considerations through the lens of segmentation

## 9.1 Introduction

### 9.1.1 The challenge

The growing volume and complexity of image data necessitate increasingly advanced analytical tools. One example of challenging tasks is image segmentation, the process of identifying and delineating structures of interest within images. Segmentation can be particularly difficult and time-consuming when dealing with large, multidimensional datasets, such as 3D volumes or time-lapse sequences, where [manual annotation](#) becomes impractical. Machine learning (ML), especially deep learning (DL), can provide effective solutions to these challenges [39].

ML algorithms learn patterns from data to perform tasks such as [image classification](#) and segmentation. Traditional ML methods, like random forest classifiers, depend on manually defined image features to classify pixels [4, 5, 12]. In contrast, DL algorithms can automatically discover and extract relevant features directly from image data using multilayer neural networks, which eliminates the need for manual feature selection. DL techniques are widely applied in complex image analysis tasks, including segmentation, [object detection](#), feature extraction, denoising, and restoration [71, 82]. Due to their ability to automatically learn hierarchical features, DL methods usually achieve greater accuracy and efficiency than traditional ML techniques [53, 88].

Segmentation greatly benefits from ML and DL, as manual segmentation is extremely time-consuming and impractical for large datasets. This chapter offers practical guidance on preparing a segmentation project and emphasises effective DL applications to tackle these challenges. While we focus on segmentation as a case study, the principles, workflows, and considerations discussed here are broadly applicable to other image analysis tasks, such as classification or denoising. Readers interested in these areas can adapt the described approaches to their specific needs.

## 9.2 Preparing for your project

### 9.2.1 Defining your task and success criteria

Every image analysis project should begin by clearly defining the scientific question you wish to answer, along with the criteria by which success will be measured. These foundational decisions will fundamentally shape your entire workflow. Careful planning of your objectives ensures that the chosen approach closely aligns with your scientific goals and will guide critical decisions about data annotation, model selection and performance evaluation.

Since segmentation serves as the central example in this chapter, it is important to first understand the different types of segmentation tasks encountered in microscopy before discussing how deep learning methods can be applied. These tasks can typically be categorised into three main types (Figure 9.1). Each segmentation type presents distinct challenges and is suited to different biological questions:

**Binary segmentation:** This is the [simplest form of segmentation](#) that separates the foreground from the background. For example, in a microscopy image, this involves distinguishing cell nuclei (foreground) from the rest of the image (background). This method is useful for detecting whether a structure is present or absent without distinguishing individual objects.

**Instance segmentation:** This [type of segmentation](#) identifies and labels each object independently. For instance, each cell in an image obtains a unique label. This method is crucial for tracking individual cells over time or measuring specific characteristics of each cell separately.

**Semantic segmentation:** This [segmentation strategy](#) involves labelling every pixel in an image according to its class, such as “nucleus,” “cytoplasm,” or “background.” Unlike instance segmentation, semantic segmentation does not differentiate between individual objects within the same class. This method is beneficial for analysing the spatial relationships and distribution of various cellular components.

Consider whether your segmentation solution is meant for a specific experiment or needs to generalise across various imaging techniques, sample types, or experimental conditions. Additionally, evaluate the volume of data to analyse, the feasibility of manual analysis, and the resources available to create a tailored image analysis pipeline. Avoid overengineering a solution when a simple analysis could provide the answer you seek.

Alongside task-specific considerations, it is equally important to clearly define the success criteria based on your objectives. For example, be prepared to answer the question, “*What do I need to accomplish for my analysis to be sufficient?*” – see [?@sec-10](#) for more information. This is important because no analysis is ever 100% accurate. Establishing these criteria early streamlines both the development and evaluation processes, ensuring that your outcomes are scientifically meaningful and practically useful (see [?@sec-10](#)).

While the following steps focus on segmentation, the underlying principles can be readily adapted to a wide range of DL tasks in microscopy.

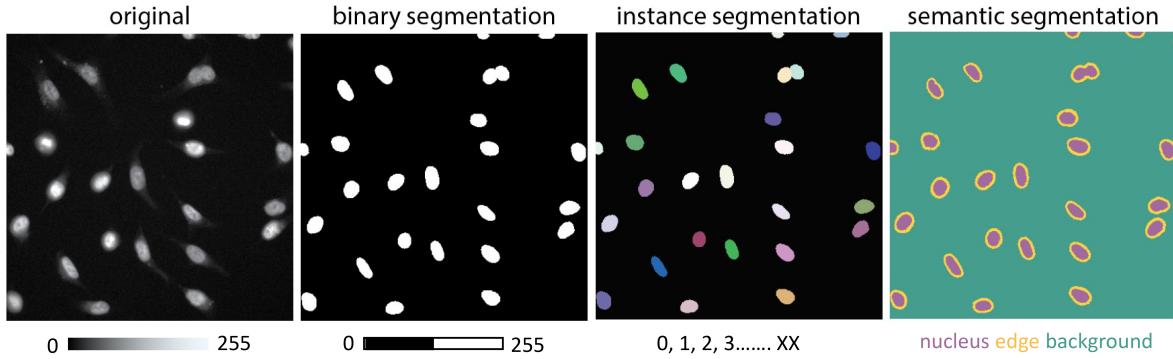


Figure 9.1: The three main types of segmentation in microscopy images. **original**: A raw grayscale fluorescence microscopy image showing cell nuclei stained with a nuclear marker. **binary segmentation**: Simplifies the image into two classes—foreground (white, nuclei) and background (black). **instance segmentation**: Assigns a unique label (shown in different colours) to each nucleus, facilitating individual object identification. **semantic segmentation**: Categorises each pixel into predefined classes—nucleus (purple), nucleus edge (yellow), and background (teal)—without distinguishing between individual objects.

### 9.2.2 Evaluating alternatives: Is DL the right choice?

Choosing the right computational method is essential for consistent and reproducible image analysis. For example in segmentation tasks, while DL can deliver exceptional segmentation performance, traditional methods and pixel classifiers still offer straightforward and efficient solutions for most tasks (Figure 9.2).

Traditional image processing techniques—such as intensity-based thresholding, morphological operations, edge detection, and filtering—are ideal for objects with clear, distinguishable features. These methods are well-documented, easy to understand, and usually require minimal computing resources. [Pixel classifiers](#), in particular, are user-friendly and can efficiently tackle many segmentation challenges with minimal manual annotation, making them highly effective for simpler analyses or smaller datasets.

DL methods excel in complex scenarios where traditional approaches fail, especially when dealing with noisy or context-dependent data. When trained on large, annotated datasets, DL models can effectively generalise across diverse imaging conditions and sample types, rapidly processing significant volumes of images. However, in the absence of pre-trained models, DL methods rarely offer shortcuts for data analysis. DL methods generally take effort and time to implement.

If you are unsure which approach to use, we usually recommend first trying classical image processing methods and pixel classifiers (Figure 9.1). We typically initiate a DL project only if these methods fail to produce satisfactory results (see Section 9.3.3.2).

## Is DL the right choice for your segmentation project?

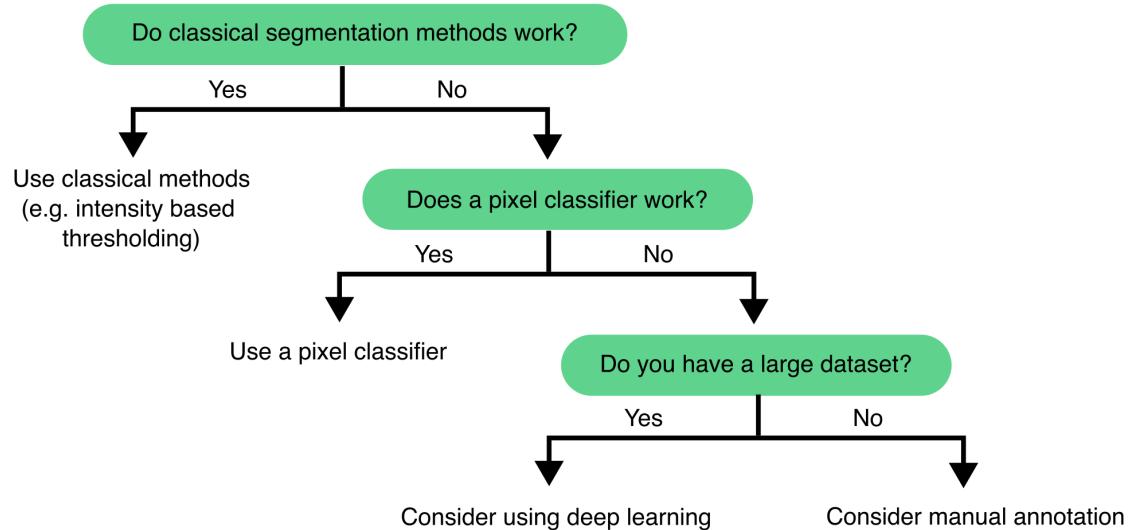


Figure 9.2: Is DL the right choice for your segmentation project? This decision tree guides the selection of appropriate segmentation approaches based on data complexity and project needs. Begin by testing classical image processing methods, such as intensity-based thresholding, which are efficient and easy to apply for well-defined features. If these methods prove insufficient, consider using a pixel classifier, which provides a user-friendly and effective solution for smaller datasets. Only consider DL if you possess a large annotated dataset and previous methods have failed. In the absence of suitable data or methods, manual annotation may be necessary.

## 9.3 Implementing a DL segmentation workflow

Although we use segmentation as our primary example, the workflow outlined in this section can be adapted to other deep learning tasks in microscopy and bioimage analysis.

### 9.3.1 Overview of a typical DL segmentation workflow

Once you decide to implement a DL approach for segmentation, the workflow can be divided into a series of steps (Figure 9.3).

The process starts by clearly defining your task (in this example, segmentation) and selecting the right DL approach (Step 1). Next, evaluate whether any existing pre-trained models can be used directly on your data or adapted for use (Step 2). If additional training is required—either from scratch or through transfer learning—prepare an appropriate training dataset that reflects your segmentation problem (Step 3). Then, train your model using the prepared dataset (Step 4) and thoroughly evaluate its performance using validation or test data (Step 5). Based on the results, you may need to refine the model by adjusting hyperparameters, improving annotations, or expanding the dataset. Once the model performs satisfactorily, it can be used to segment new, unseen data (Step 6). We will now discuss each step in more detail.

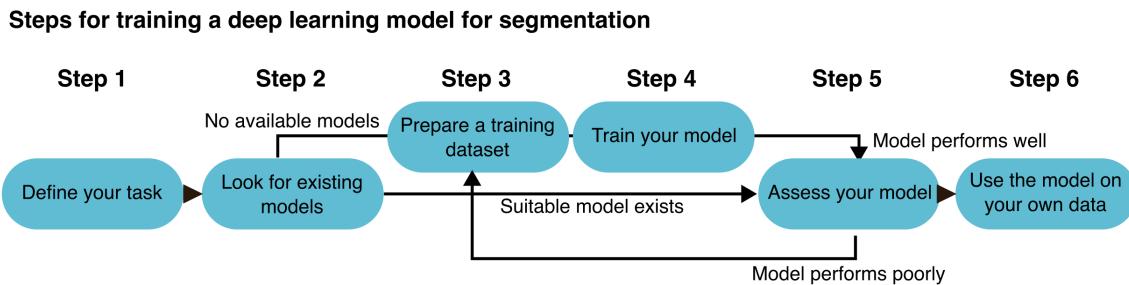


Figure 9.3: Conceptual workflow for training a DL segmentation model. The workflow begins with defining the segmentation task (Step 1), followed by searching for suitable pre-trained models (Step 2). If no such model exists, a training dataset must be prepared (Step 3), and the model is trained (Step 4). The trained model is then evaluated on validation or test data (Step 5). If it performs well, it can be applied to new data (Step 6); otherwise, the model is iteratively refined by returning to earlier steps. Feedback loops from evaluation to earlier stages help refine and improve the model accuracy.

### **9.3.2 Selecting a suitable DL approach**

The first step in choosing a DL approach for image segmentation is to clearly define your segmentation task, whether it's binary, semantic, or instance segmentation (Figure 9.1), and to determine if you require 2D or 3D segmentation. Next, you should consider whether the model or tool you plan to use makes assumptions about the shapes or structures of the objects you want to segment. Understanding these assumptions will aid in selecting a model that fits your specific biological problem (see Section 9.2.1). Additionally, consider the amount of data that needs annotation for a particular DL approach (see Section 9.3.4.5). Finally, take into account your available computational resources (see Section 9.4.2). More complex models typically demand more GPU memory, longer training times, and additional storage, especially for 3D data or large datasets.

For example, StarDist [100], a widely used tool for nuclei segmentation, assumes that objects are **star-convex polygons**: i.e., a shape for which any two points on its boundary can be connected by a single line that does not intersect the boundary. This assumption works well for round or oval shapes but makes StarDist less suitable for segmenting irregularly shaped or elongated structures. In contrast, Cellpose [107] uses spatial vector flows to direct pixels toward object centres. This approach enables Cellpose to segment objects of various shapes and sizes, including irregular, elongated, or non-convex forms.

Choosing the right DL strategy requires aligning your goal, object shape, data dimensionality, and computing capacity with the strengths and assumptions of the available DL architectures.

### **9.3.3 Deciding whether to train a new model**

#### **9.3.3.1 Leveraging pre-trained models**

The increasing availability of already trained (pre-trained) DL models has greatly simplified image analysis. Many of these models can be directly applied to your data, removing the need to train your own model [11, 28]. This reduces the technical barrier and saves time, making advanced analysis more accessible. However, it is essential to evaluate the quality of any pre-trained model before relying on its results (Figure 9.4). A model that performs well in one context may not be as effective on your specific data. Always conduct quality control by visually inspecting the outputs and assessing performance with quantitative metrics such as Intersection over Union (IoU) or F1-score, using a small, representative test set. This step is vital when model predictions are used in downstream analyses (see Section 9.3.7).

Another significant benefit of pre-trained models is their adaptability. Instead of starting from scratch, you can often fine-tune an existing model (see Section 9.3.6). This method entails retraining the model with a smaller, task-specific dataset, enabling it to adjust to your images while requiring far fewer annotations.

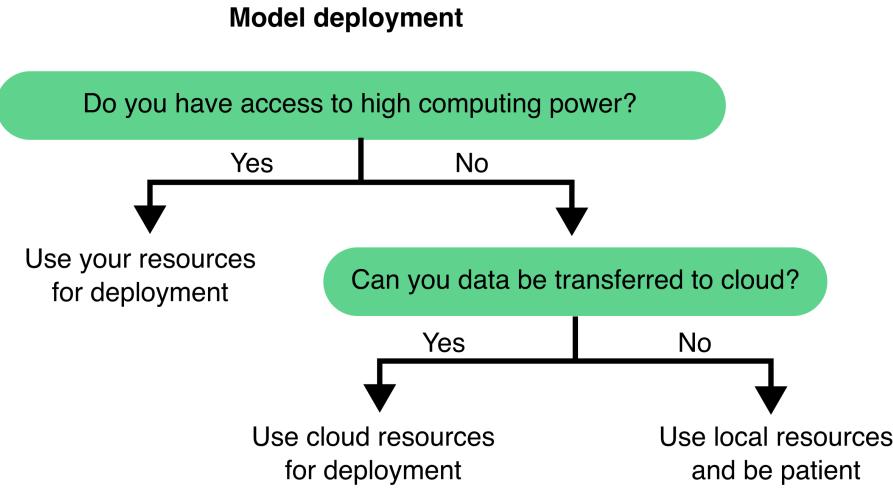


Figure 9.4: Decision workflow for selecting a DL model training approach. This flowchart outlines how to determine an appropriate training approach based on the availability and performance of pre-trained models.

Several excellent resources host pre-trained models suitable for microscopy. Researchers also increasingly share trained models alongside their datasets and publications, promoting open science. Platforms like Zenodo are commonly used for this purpose [26, 29], although deployment may require handling specific file formats or environments (see Chapter 8 for more information).

### 9.3.3.2 When to train your model

Pre-trained models serve as an excellent starting point for various microscopy tasks. However, there are many scenarios where training a custom model becomes essential. Custom training enables the model to learn the specific characteristics of your dataset, experiment, or imaging modality, resulting in enhanced performance [3, 79, 16]. This is particularly crucial when your data differs significantly from the data used to train existing models. Thus, their performance should always be validated. If quality assessment metrics are poor or key features are not accurately segmented, consider training your own model.

Ultimately, always evaluate the model's performance against your defined success criteria (see Section 9.3.7). Custom training may be the best path forward if the current model does not meet your needs.

### 9.3.4 Preparing your dataset for training

A well-designed training dataset is essential for developing successful DL models on tasks such as segmentation. The number of images and the quality of annotations needed vary based on factors such as task complexity and the architecture of the intended model.

#### 9.3.4.1 Types of model training

For segmentation, most DL models are trained using supervised learning, where each input image is paired with a manually annotated ground truth mask. In this context, all objects that need segmentation must be annotated in the training dataset. This approach enables the model to learn a direct mapping from raw images to segmentation outputs (Figure 9.7).

However, alternative approaches can help reduce the need for extensive manual annotations:

- **Unsupervised learning** trains models without paired input and output data. Instead, the network identifies patterns or similarities in unlabelled images [51].
- **Self-supervised learning** involves designing tasks in which the model learns useful features directly from the input data without needing explicit labels [63].
- **Weakly supervised learning** uses partial, noisy, or imprecise labels to guide training, which can significantly reduce annotation effort [15, 73].

#### 9.3.4.2 Creating Manual Annotations

Creating accurate annotations manually is time-consuming, particularly for 3D datasets. Tools like Fiji [98], Napari [1], and QuPath [7] are frequently employed for manual labelling. Typically, manual annotation involves drawing each object on the image and converting it into a mask or label.

##### 💡 Here it is an example pipeline for manually annotating data using Fiji

See also [26].

###### 1. Open Fiji – activate the LOCI update site and restart Fiji.

LOCI tools are required for exporting ROI maps. To enable them, go to *Help > Update > Manage Update Sites*, look for ‘LOCI’ and check the *Active* checkbox. Then, click on *Apply and Close* and *Apply Changes*, this update site ensures the necessary plugins are installed. Finally, restart Fiji.

###### 2. Open your image you wish to annotate.

Use *File > Open* to browse and load the microscopy image that you want to label

manually. You can also drag and drop your image to Fiji.

**3. Select the Oval or Freehand selection tool.**

These tools, found in the Fiji toolbar, allow you to manually outline the structures of interest in your image.

**4. Start drawing around each object (yes, each one!).**

Carefully trace each cell or feature you want to annotate—precision is key to ensure useful training data for DL.

**5. After drawing each object, press “t” on your keyboard → the selection will be stored in the ROI manager.**

This adds the drawn region to the ROI (Region of Interest) list, keeping track of all annotated objects in the image.

**6. Repeat until all objects are in the ROI manager.**

Continue drawing and pressing “t” until you have annotated every relevant object in the image.

**7. When finished, go to *Plugins* › *LOCI* › *ROI Map*.**

This plugin converts all saved ROIs into a single labeled ROI map image, assigning unique values to each region.

**8. Save the generated ROI map with the same title as the original image in a separate folder.**

Consistent naming ensures each annotated map can be correctly matched with its corresponding raw image during training or analysis.

**9. At the end, you will have one folder with the original images and another for the ROI maps.**

This separation makes it easier to organise and use your data with image analysis or DL pipelines.

#### **9.3.4.3 Accelerating annotation with automatic initial segmentations**

Creating high-quality annotations often represents the most time-consuming aspect of training a DL model for segmentation. To alleviate this burden, you can start from automatically produced initial segmentations. For example, using simple thresholding methods such as Otsu’s thresholding to generate rough segmentations can decrease the total annotation time. Even more powerfully, pre-trained DL models such as those provided with StarDist [100] and Cellpose [107] can generate more accurate initial segmentations that users can manually refine. These annotations can then be used to retrain the model, establishing an iterative cycle that accelerates both labelling and model refinement.

New tools are also pushing the boundaries of interactive annotation. For example, Segment Anything for Microscopy (SAM) [3] facilitates automatic and user-guided segmentation and allows the model to be retrained on user-provided data. Similarly, Cellpose 2.0 [79] features a human-in-the-loop workflow, allowing users to edit DL-generated segmentations. This hybrid approach enhances accuracy while significantly reducing the time and effort required for manual annotation.

#### 9.3.4.4 Expanding your dataset with augmentation and synthetic data

When the number of training samples is limited, [augmentation](#) techniques can enhance dataset diversity to improve the model’s generalisation ability and performance on validation and testing [56, 103]. Common augmentation strategies include image rotation, flipping, scaling, and contrast adjustment. However, it’s important to apply augmentation carefully, as excessive or unrealistic augmentation can confuse the model or cause it to learn patterns that do not exist in real data.

In the absence of sufficient real data, synthetic data generated through simulations or [domain randomization](#) can help pre-training a model [62, 86]. These synthetic samples can expose the model to a broader range of scenarios early in training, before transitioning to fine-tuning with real, annotated data.

In summary, a successful segmentation pipeline relies on a careful balance between data quantity and annotation quality. Augmentation strategies can efficiently help to scale and balance training datasets.

#### 9.3.4.5 Choosing the dataset size: specific vs. general models

In supervised training, it is crucial that each image in the training set is accompanied by a corresponding label image (see Section 9.3.4.1). The number of image-label pairs required depends on the number of labels per image, the complexity of the model and the desired level of generalisability. Still, the key is having enough representative examples and corresponding annotations for the model to learn meaningful patterns.

Small and well-curated datasets consisting of tens of images may suffice for highly specific applications, such as segmenting cells or nuclei using a defined imaging modality [16]. In these scenarios, [transfer learning](#) can also be especially beneficial (see Section 9.3.6). Models designed to generalise across a wide range of conditions, tissue types, or imaging modalities typically require much larger and more diverse datasets (hundreds to thousands of annotated images) [107]. These datasets are essential for capturing the inherent variability in broader use cases.

### 9.3.5 Training a segmentation model from scratch

Once you have annotated your training dataset, the next steps are to organise your data for training, initialise your model by selecting appropriate [hyperparameters](#), and start the training process (Figure 9.7).

#### 9.3.5.1 Splitting your training data: training, validation, and test sets

A crucial part of preparing your dataset is dividing it into three subsets: training, validation, and test sets. Each subset should contain the original microscopy images paired with their corresponding ground truth segmentations. A common strategy is to allocate 70–80% of the data for training, 10–15% for validation, and the remainder for testing. To ensure unbiased evaluation, ensure these subsets do not overlap in terms of fields of view, represent the variability of your entire dataset, and are randomly assigned to each set respectively.

The **training set** is used to train the model to recognise relevant features. To enhance generalisation, it must encompass a broad spectrum of scenarios and image conditions. Otherwise, the model risks overfitting—excelling with the training data but faltering with new images (Figure 9.6).

The **validation set** is used during training to provide feedback on the model’s performance with unseen data. This feedback, conveyed as validation loss, assists in detecting overfitting (Figure 9.6), guiding hyperparameter tuning (see Section 9.3.5.3), and informing training decisions. Although a separate validation set is ideal, many workflows create one in practice by reserving a portion (typically 10% to 30%) of the training data.

The **test set**, which serves a separate role, evaluates the model’s performance on entirely unseen data. Unlike the validation set, the test set is not utilised during training, ensuring an unbiased performance assessment. Test images should also include ground truth annotations to facilitate quantitative quality control. Reporting test set performance, using metrics such as accuracy, IoU, or F1-score, is crucial, especially when publishing or benchmarking your model [55].

#### 9.3.5.2 Understanding the training process

A DL model is composed of multiple layers (Figure 9.5). Each layer contains tens to hundreds of image processing operations (typically multiplications or convolutions), each controlled by multiple adjustable parameters (called weights). Altogether, a DL model may contain millions of adjustable weights. When an input image is processed by a DL model, it is sequentially processed by each layer until an output is generated. Segmentation tasks typically involve converting input images into labelled outputs. During training, the model weights are modified as the model learns how to perform a specific task.

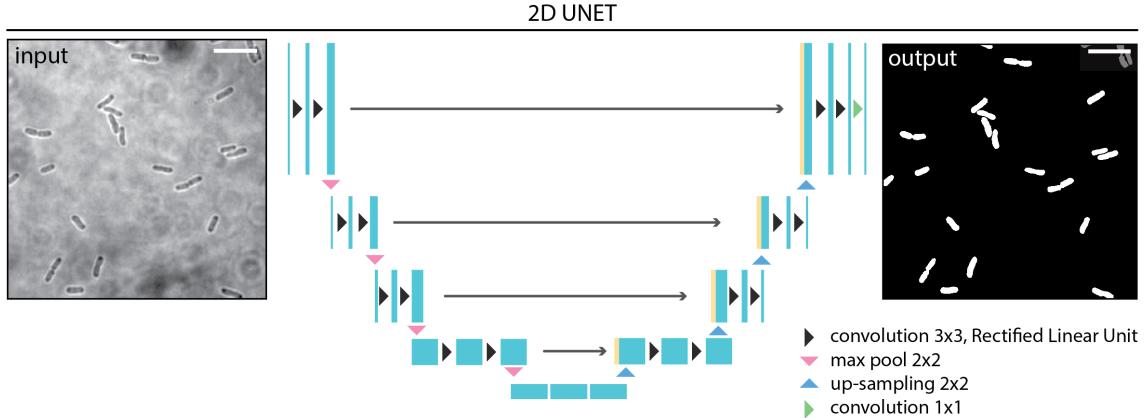


Figure 9.5: 2D U-Net architecture for image segmentation. It applies layers of convolutions, pooling, and upsampling to extract features and generate labelled segmentation masks. During training, model weights are iteratively adjusted based on the difference between predictions and ground truth labels, using a loss function and backpropagation.

Training begins with initialising these weights. When training from scratch, the initialisation is often random. However, when using a pre-trained model, the weights are already optimized based on previous training, providing the model with a significant head start (see Section 9.3.6).

The training process is iterative (Figure 9.7). Each cycle of training is called an [epoch](#). During each epoch, the model typically learns from every image in the training set. Since datasets are often too large to fit into memory all at once, each epoch is divided into steps or iterations, with each step processing a smaller subset of the data known as a batch. The batch size determines how many samples are processed simultaneously.

During each step, the model generates predictions for the current data batch. These predictions are compared to the ground truth labels using a loss function that calculates the similarity between the predictions and the ground truths. This score is called the training loss. The model utilises this feedback to adjust its weights through a process known as [backpropagation](#), guided by an optimisation algorithm, to improve its accuracy in future iterations.

At the end of each epoch, the model assesses its performance on the validation set, which comprises data it has not encountered during training. This produces the validation loss, indicating how well the model generalises to new data.

Monitoring both training and validation losses during training helps determine whether the model is learning effectively. A consistently decreasing validation loss indicates that the model is improving and generalising well (see Section 9.3.5.4).

### 9.3.5.3 Choosing your model hyperparameters

Now that you understand the training process, the next step is to configure the model’s hyperparameters, which are the settings that dictate how the model learns. While the model’s parameters (its weights) are updated during training, hyperparameters are established beforehand, defining the structure and behaviour of the training process. Below are some of the most common hyperparameters and their effects on training:

- **Batch size:** This refers to the number of images processed simultaneously in each training step. Smaller batch sizes are less demanding on memory and may enhance generalisation, although they can result in slower training. In contrast, larger batch sizes accelerate training but necessitate more GPU memory.
- **Epochs:** An epoch refers to a training cycle in which the model processes the entire training dataset. Increasing the number of epochs allows the model to learn more, but also raises the risk of overfitting. More is not always better; it is essential to monitor performance on the validation set.
- **Learning rate:** It determines the extent to which the model’s weights are adjusted during training. A high learning rate can result in quicker training but may overshoot the optimal solution. Conversely, a low learning rate provides more stable progress, although it may slow down convergence.
- **Optimizer:** An algorithm that updates weights to minimise the loss function. Common optimisers include SGD (stochastic gradient descent) and Adam (adaptive moment estimation), the latter being widely used for its adaptive learning rate and robust performance.
- **Learning rate scheduler:** Dynamically adjusts the learning rate during training, typically decreasing it after a specific number of epochs or when the validation loss plateaus. This approach helps balance rapid early learning with more refined convergence later on.
- **Patch size:** Instead of using full-resolution images, smaller patches are often utilised for training to reduce memory usage and enhance training speed. The patch size is determined by both available resources and the scale of the structures to be segmented.
- **Patience (early stopping):** This parameter defines the number of epochs to wait before halting training if the validation loss does not improve. It helps prevent wasting resources on overfitting and overtraining.

Given the many possible configurations, tuning hyperparameters is often essential—especially when applying a model to new data. Start with the recommended values from the model’s original publication, but you might need to conduct a hyperparameter search to optimize performance. This can range from a simple grid search to more advanced methods, such as Gaussian process-based Bayesian optimisation [44] or genetic algorithms [2].

#### 9.3.5.4 Monitoring training and validation Losses

Once your model begins training, it is helpful to evaluate its learning progress. The two key metrics for assessment are the training loss and the validation loss (Figure 9.6). Monitoring both throughout the training process offers insight into whether your model is improving and learning to generalise beyond the training data. The three main behaviours that you may encounter during training are:



Figure 9.6: Monitoring training and validation losses during model training. The plot illustrates three typical learning behaviours: **underfitting** (both losses remain high and similar), **good fitting** (both losses decrease, with validation loss slightly higher), and **overfitting** (training loss continues to decrease while validation loss plateaus or increases), highlighting the importance of tracking these metrics to assess model performance and generalisation.

- **Underfitting:** The model has been trained with insufficient data or for too few epochs, resulting in similar training and validation losses, which is far from optimal.
- **Good fitting:** Both training and validation losses decrease, with the validation loss slightly higher (worse) than the training loss, which is expected. This represents the ideal scenario.
- **Overfitting:** The model achieves an excellent training loss, but the validation loss does not improve and may in fact diverge. This may indicate overly similar training data or excessive training epochs, preventing the model from generalising to new data.

#### 9.3.6 Fine-Tuning Pre-existing Models

Instead of training a model from scratch, fine-tuning an existing DL model is usually more efficient, especially when your data resembles the dataset used to train the original model. This

approach utilises pre-trained weights and previously learned features, significantly decreasing the amount of required annotated data, training time, and computational resources.

#### **9.3.6.1 Applying Transfer Learning**

Transfer learning refers to the process of taking a pre-trained model and adapting it to a new but related task by providing task-specific training data, typically in the form of manually annotated image pairs (Figure 9.7). Transfer learning typically involves freezing part of the model (for instance, the initial layers or all layers except the last ones), so their weights are not updated during training. Only the unfrozen layers are updated when the model is trained on new data. Then, the model is trained on the new data, but only the layers that you have unfrozen will be updated. Since the base model already encodes many useful low-level features (e.g., edges, shapes, textures), this approach allows researchers to focus on refining the model for their specific biological structures or imaging modalities [61, 72].

This method is especially effective when:

- You have limited training data available.
- Your imaging conditions closely match those of the pre-trained model.
- You wish to quickly adapt a general model to a specific dataset.

#### **9.3.6.2 Conducting fine-tuning**

In classic fine-tuning, all layers of the pre-trained model are retrained, with their weights initialised from the original training (Figure 9.7) [3, 79]. Thus, you continue training the full model using the new data. This approach allows the model to adjust more comprehensively to new data while still preserving the advantages of pre-learned features.

Classic fine-tuning is ideal when:

- Your dataset is moderately different from the original training data (e.g., the same biological structure but different staining or modality).
- You expect that earlier layers may need to adapt, not just the final classifier or output layers. Early layers in deep networks typically learn to detect general features such as edges, textures, or simple shapes, while later layers capture more complex, task-specific patterns. If your new data differs in basic appearance or imaging modality, updating the early layers helps the model better extract relevant low-level features from your images.

- You have enough annotated data to avoid overfitting during full model training. Although this method is more computationally demanding than transfer learning, where only a subset of layers are retrained, it often leads to better results on diverse datasets..

### **9.3.6.3 Iterative training: keeping humans in the loop**

Iterative fine-tuning is an interactive approach that combines model prediction with human annotation (see Section 9.3.4.3). The workflow typically starts with a pre-trained model predicting new images. A user then manually corrects or annotates these predictions, and the improved annotations update the model (Figure 9.7). This cycle continues, progressively enhancing the model’s accuracy with each iteration until it performs as expected [20, 23, 79].

This method is particularly powerful when:

- Annotated data is scarce or expensive to generate.
- You work with rare structures, unusual imaging conditions, or new experimental systems.
- You want to efficiently build a custom model using feedback from domain experts.

### **9.3.7 Evaluating the performance of your model**

With the rapid increase in DL tools and pre-trained models, it has become easier to use DL for image segmentation, but harder to determine which model will work best for your data. Regardless of how promising a model appears, you must always evaluate its performance before trusting its results. Evaluation is not optional; it is a critical step to ensure that the model meets the requirements of your specific task [55] (Figure 9.8).

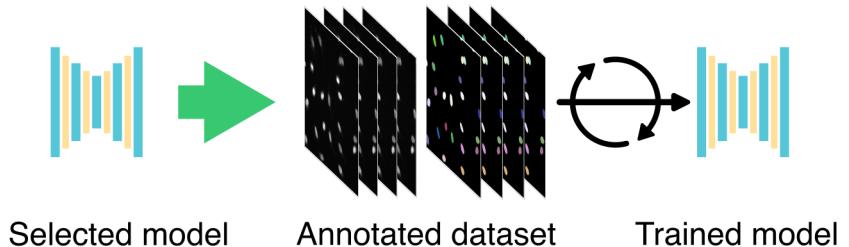
There are two main ways to evaluate a model:

- Qualitative evaluation entails visually inspecting the model’s predictions. This approach can help you quickly identify clear errors or failures. It is effective for a small number of images, but it becomes impractical for large datasets or for comparing similar-looking outputs across multiple models.
- Quantitative evaluation provides objective metrics for comparing models and tracking improvements. To achieve this, you need a small, labelled test set (typically 5 to 10 images with accurate ground truth segmentations). This test set must remain independent of your training and validation data to ensure an unbiased assessment.

Common metrics used in quantitative evaluation include:

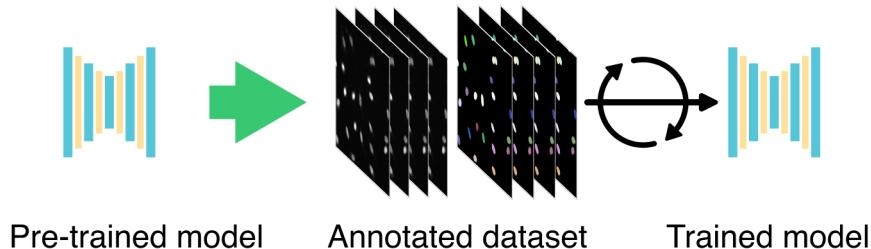
### Training from Scratch

A model is trained from randomly initialized weights using a large annotation dataset



### Transfer Learning or Fine-Tuning

A pre-trained model is adapted to a new task by adding new annotated data



### Human-in-the-Loop Fine-Tuning

Model predictions are iteratively corrected through manual annotations

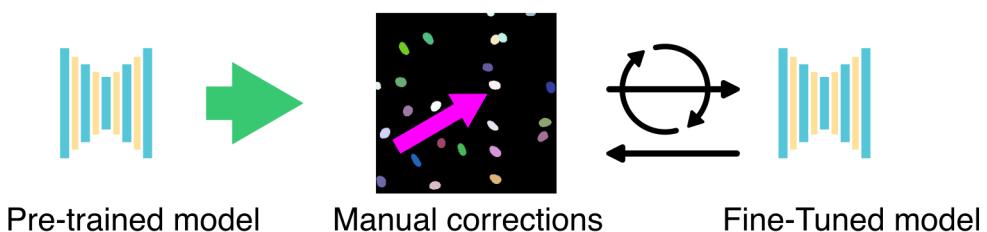


Figure 9.7: Strategies for training DL models for image segmentation. A model can be **trained from scratch** using a large annotated dataset, **fine-tuned** from a pre-trained model with task-specific data, or refined through **human-in-the-loop** workflows where model predictions are manually corrected and fed back for retraining. These approaches balance performance, data availability, and annotation effort.

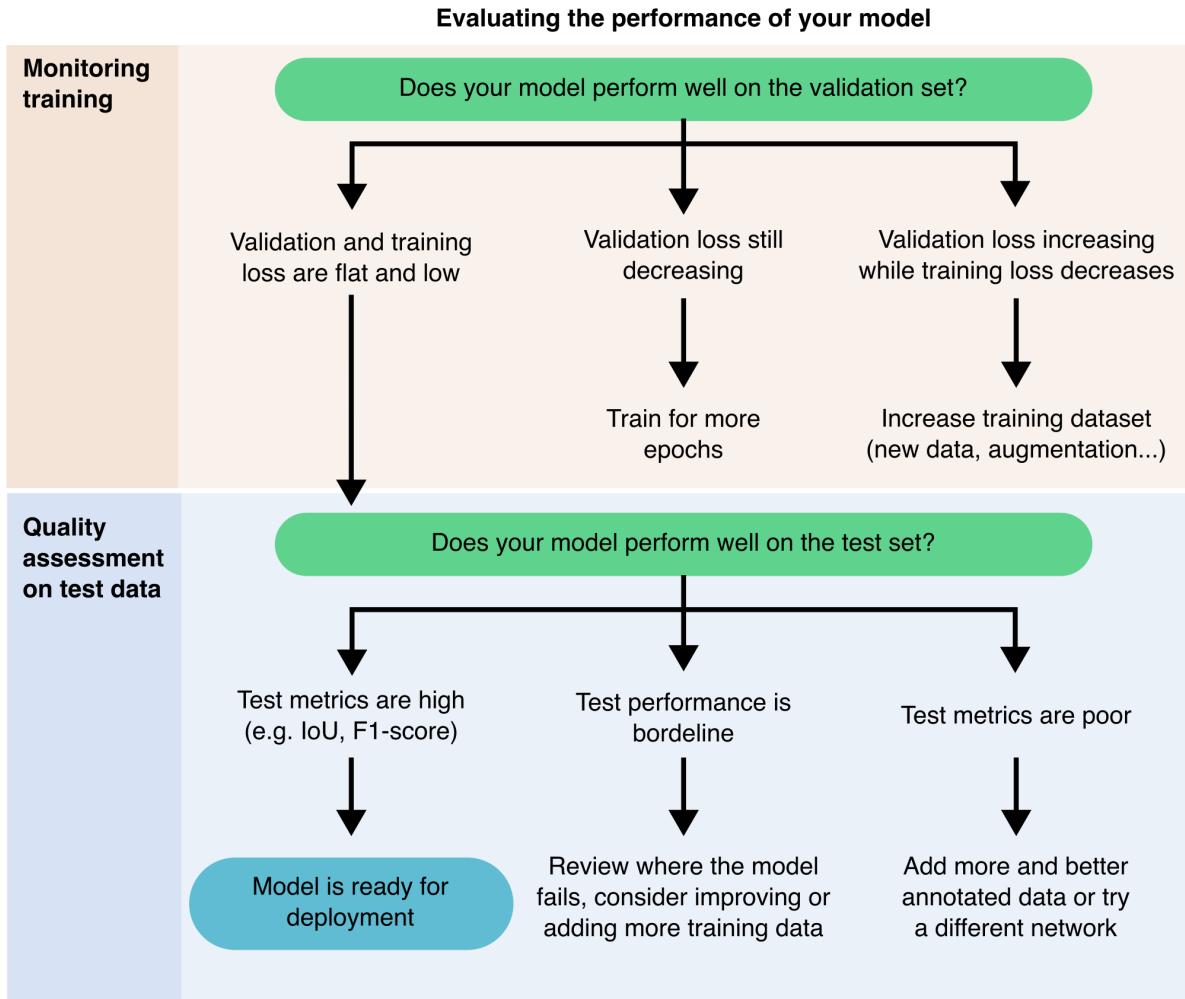


Figure 9.8: Workflow for evaluating DL model performance during training and on test data. Evaluating model performance is essential before deploying any DL model for image segmentation. This diagram outlines a two-stage process: assessment during training and on a separate test set. During training, validation and training losses (see Section 9.3.5.4) guide whether to continue training, stop, or expand the dataset. After training, performance is evaluated using a test set. High test metrics (e.g., IoU, F1-score) indicate readiness for deployment. Borderline or poor results suggest reviewing errors, refining training data, or trying a different model. This approach ensures model reliability and task-specific performance.

- Intersection over Union (IoU), also known as the Jaccard Index, measures the overlap between the predicted segmentation and the ground truth.
- F1-score (Dice coefficient): This is especially valuable when the object of interest covers a small area in the image, as it balances precision and recall.
- **True Positives** (TP), **False Positives** (FP), and **False Negatives** (FN) are particularly important in semantic segmentation and can be used to calculate the **IoU** or **F1 score**.

For more information on these metrics, we recommend [55] (also see [?@sec-10](#) for more information).

If a model fails to produce reasonable results, even on simple examples, you can often reject it based solely on qualitative inspection. However, in these cases, quantitative metrics can still help you understand how and where the model fails.

If your evaluation metrics indicate weak performance, especially for certain structures or image types, you may need to fine-tune the model (see Section [9.3.6](#)). Consistently strong scores across various test images suggest that a model could be dependable and ready for deployment. If no pre-trained model meets your expectations, the best course may be to train your model using your images (see Section [9.3.5](#)).

In summary, **never skip evaluation**. Every model must be tested—both visually and quantitatively—to ensure it truly works for your data and provides results you can trust.

### **9.3.8 Deploying your model on new data**

Once a segmentation model has been trained and validated, it can be used on new, unseen images. This step typically involves feeding new images into the model to generate segmentation predictions. The deployment approach relies on the computational resources (see Section [9.4.2](#)) as well as the size and complexity of your dataset (Figure [9.9](#)).

### **9.3.9 Troubleshooting Common Problems**

**I found a tool or DL model online, but it does not work. What should I do?**

**When should I train a model or segment manually?**

Refer to (Section [9.3.3.2](#)) for more details, but generally, this decision depends on your dataset and the performance of existing pre-trained models (Figure [9.10](#)). If you only need to segment a small number of images, manually segmenting them is often the quickest and simplest solution. However, if you are dealing with a large dataset, it may be more efficient to annotate a small subset and use it to train a deep-learning model that can automate the segmentation of the rest.

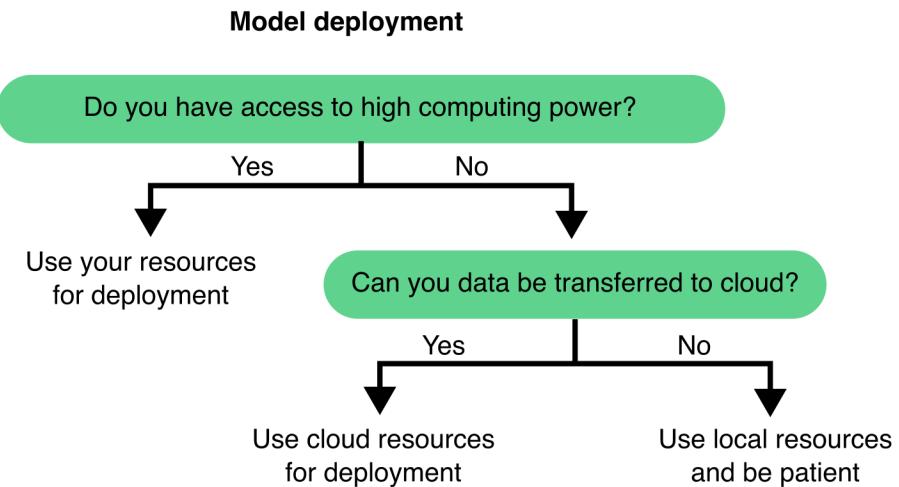


Figure 9.9: Decision workflow for model deployment strategy based on computational resources. The choice of deployment strategy depends on the availability of computational resources (see Section 9.4.2) and the sensitivity of the data . If high-performance computing resources are available locally, these should be used for deployment. In their absence, consider whether the data can be transferred to the cloud. If so, cloud-based resources offer an efficient solution. However, if data transfer is restricted—due to size or sensitivity—local deployment remains the only option, though it may require significantly more time.

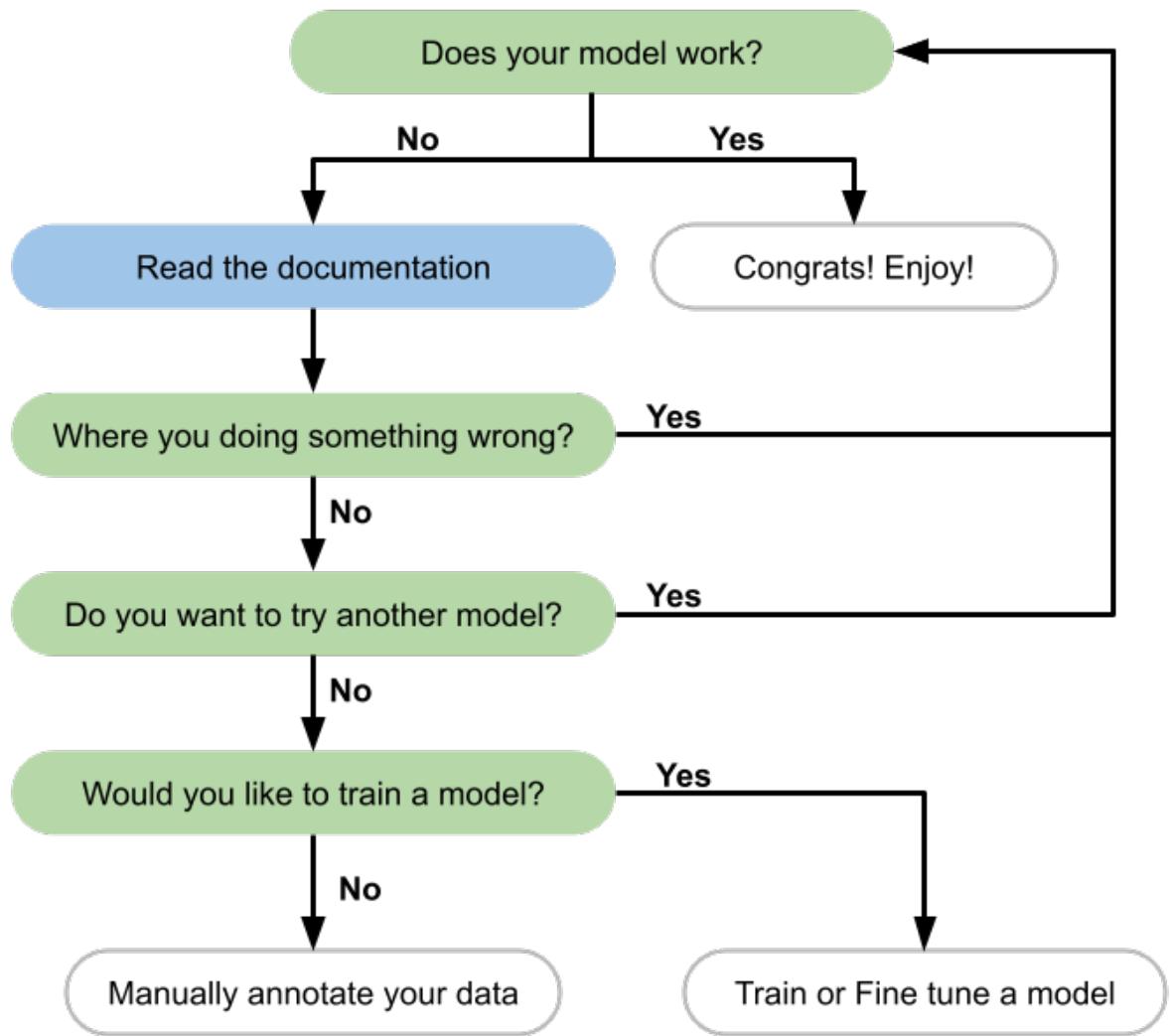


Figure 9.10: Common DL segmentation problems and troubleshooting tips.

## I decided to train my DL model, but it is not performing correctly. What should I do?

First, ensure that you have trained the model for a sufficient number of epochs—this depends on the size of your dataset and the architecture of the model. Check the training loss: if it has plateaued, your model may be fully trained. If it is still decreasing, continue training.

If training is completed but results are poor, examine your data. Is the model missing specific features? Are there types of cells or structures that it consistently fails to segment? If so, ensure those examples are well represented and correctly annotated in your training data. You may need to enhance or expand your annotations.

If performance is poor, you may need additional annotated data to help the model generalise more effectively (Figure 9.8). Consider the following questions:

- Is my dataset balanced? Does it include sufficient examples of each structure or class I want to segment?
- Am I training on one experimental batch while validating or testing on another?

## How many images should I have to train my model?

Refer to Section 9.2.1 for more details. There's no one-size-fits-all answer—it depends on the complexity of your task, your model architecture, and the variability in your data. More complex tasks typically require more data. Larger images can also be broken into more patches, effectively increasing your dataset size. While few-shot models are being developed for small datasets, most established DL models require a substantial amount of data.

## Possible technical issues that you may encounter when training your DL model.

- The model predicts the same class for all pixels or segments in every cell. Your dataset might be unbalanced, containing too many similar examples. Adding more diverse or underrepresented examples can help the model learn to differentiate between classes.
- Out-of-memory errors during training: Consider reducing the batch size or the image patch size. If that doesn't resolve the issue, consider switching to a workstation or cloud service with greater computational capacity.
- The model performs well on training data but poorly on new images, suggesting overfitting (Figure 9.6). Implement data augmentation and increase dataset diversity to help the model generalise better.
- Inconsistent results across different computers: Differences in GPUs or environments can cause slight variations in outcomes. If the differences are significant, verify that all systems use consistent software versions and configurations. For further information on this topic, refer to Section 9.4.3.

## 9.4 Further considerations for DL segmentation

### 9.4.1 Choosing the Right Tools for DL

Selecting the right tools to train and use DL models depends mainly on your level of programming experience and comfort with technical interfaces.

If you prefer not to write code or use command-line tools, opt for platforms that offer graphical user interfaces (GUIs) or interactive notebooks with pre-configured workflows. These tools let you perform powerful segmentation tasks using intuitive interfaces and simple widgets.

GUI-based tools include, for instance (see Chapter 8 for more tools):

- Cellpose GUI
- Fiji with DeepImageJ and StarDist plugins
- Napari
- Ilastik
- QuPath

Interactive Jupyter notebooks provide a flexible balance between code and GUI. They enable you to execute code in manageable steps (cells) and immediately see the results. Tools like BiaPy, and DL4MicEverywhere [41] leverage Jupyter notebooks, concealing complex code behind user-friendly interfaces. These platforms cater to users with little or no coding experience while still allowing advanced users to access and modify code as needed. DL4MicEverywhere, in particular, established a widely adopted framework for training DL models via notebooks, contributing to the standardisation and simplification of the workflow.

If you are comfortable with programming, you will have even more flexibility. Languages such as Python, MATLAB, Julia, Java, and Rust provide options for building and customizing DL workflows. Python stands out as the most beginner-friendly and widely supported choice, boasting a large ecosystem of libraries and community support. Popular Python libraries for DL include PyTorch, TensorFlow, Keras, and JAX.

While coding can involve a steeper learning curve, it allows you to create customized pipelines, integrate various tools, and troubleshoot intricate workflows, unlocking the full potential of DL for microscopy segmentation.

#### 9.4.2 Managing Computational Resources

When using DL for microscopy, an important consideration is the availability and capacity of your computational resources (Figure 9.9). High-performance DL models, particularly those used for 3D image data, can be very demanding regarding memory and processing power.

When selecting or designing a DL model, evaluate your available infrastructure:

- GPU memory: Determines how large your model and batch size can be.
- Training time: Influences your ability to iterate quickly; simpler models train faster.
- Dataset size: Larger datasets benefit from more powerful hardware and longer training times.

A practical strategy involves starting with lightweight models that demand fewer resources and scaling up to more complex architectures only if performance improvements become necessary. Tools like StarDist and Cellpose, for example, provide efficient options that function effectively with relatively modest hardware.

Additionally, consider whether to train and deploy your model locally or in the cloud (Figure 9.11). Local training is often more feasible if you already have access to a compatible workstation and want full control over data and execution. However, cloud-based services like Google Colab or AWS offer access to more powerful hardware, removing the need for local infrastructure—this is especially beneficial when working with large models or 3D datasets.

There are four typical combinations of training and prediction workflows:

- Training and prediction locally is well suited for small to medium-sized datasets, especially when computational demands are moderate and data privacy is a priority. This approach also supports some user-friendly desktop applications, such as the Cellpose 2.0 [79], which can be run locally without requiring cloud access or advanced technical setup.
- Training locally, prediction in the cloud may be useful when models are trained in-house but need to be deployed at scale or integrated into cloud-based pipelines.
- Training in the cloud, prediction locally enables researchers to take advantage of powerful cloud GPUs for model development, while keeping inference close to the data source (e.g., a microscope workstation or in case of sensitive data).
- Training and prediction in the cloud is well suited for collaborative projects or large-scale deployments, where access to centralized, scalable infrastructure is critical.

Choosing between these strategies depends on your data size, hardware access, choice of software, collaboration needs, and whether your workflow prioritizes flexibility, scalability, or control.

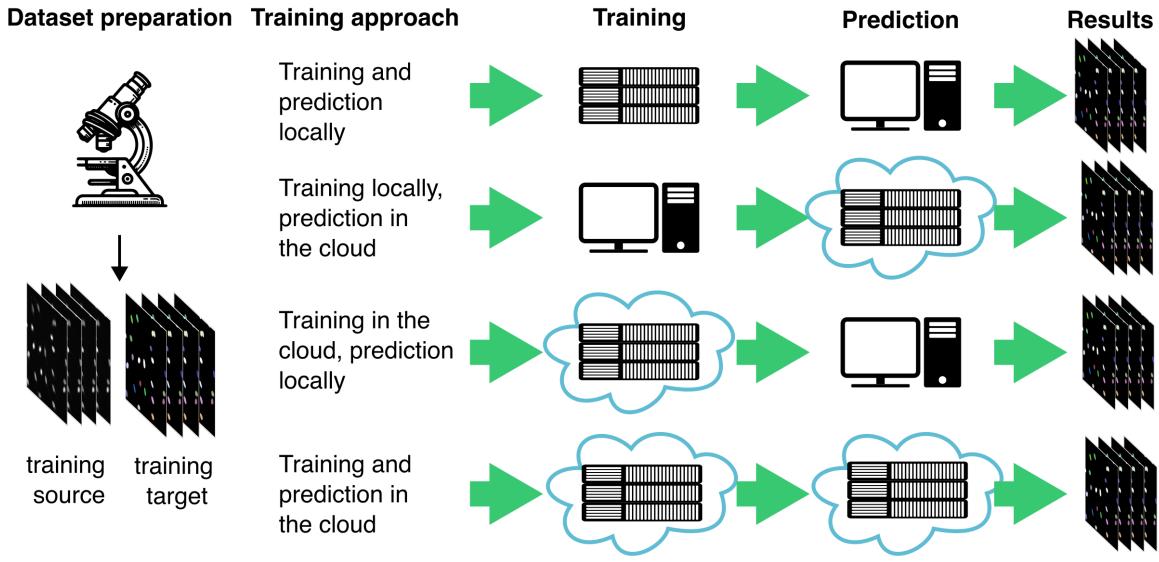


Figure 9.11: Training and deployment strategies for DL models in microscopy. Depending on the available tools and infrastructure, models can be trained and deployed locally or in the cloud. Modified from [16].

#### 9.4.3 Ensuring Reproducibility in DL

When sharing how you trained a DL model, two key elements often come to mind: the dataset used and the code that runs the model. However, in practice, reproducibility extends beyond just data and code. In programming environments like Python, which rely heavily on external libraries, ensuring reproducibility also requires capturing the exact configuration of the environment in which the model was trained.

DL models are sensitive to changes in library versions and dependencies. Even minor differences in the software stack can result in inconsistent outcomes or training failures. While sharing a list of dependencies (e.g., a requirements.txt or a Conda environment file) is a constructive step, differences in operating systems or local setups can still lead to issues.

A robust and increasingly popular solution is containerisation. Containers package software, dependencies, and environment settings into a portable and self-contained unit. One of the most widely used containerization tools is Docker. A Docker container can be considered a lightweight, standalone [virtual machine](#) that includes everything needed to run code, such as the operating system, libraries, and runtime, ensuring applications run consistently across different machines.

Using containers ensures that your model training and inference processes remain consistent, no matter who executes them or where they are conducted. This greatly simplifies the ability

of collaborators or reviewers to reproduce your results.

For researchers unfamiliar with software development, tools like DL4MicEverywhere [41] and bia-binder [90] simplify the use of containers by integrating them into user-friendly Jupyter notebook environments. These platforms enable researchers to benefit from the reproducibility of containers without needing to manage complex setups or command-line tools.

Reproducibility is crucial for establishing trust in computational results and facilitating long-term scientific collaboration. To ensure your DL workflows are reproducible, follow these best practices:

- Pin every software version used in your workflow.
- Document your environment setup thoroughly.
- Provide a containerised version of your training and inference pipeline when possible.
- Taking these steps will make it easier for others to reproduce your results, build on your work, and apply your models in different research settings.

For more information on best practices, consult [55].

## 9.5 Summary & Outlook

Segmenting microscopy images remains a critical yet challenging task in bioimage analysis. In this chapter, we have used segmentation as a representative example to illustrate deep learning workflows and considerations. However, the strategies and best practices described here—such as data preparation, model selection, training, evaluation, and deployment—are relevant to a wide range of image analysis tasks, including classification, detection, and tracking. DL has undeniably transformed this field, offering robust solutions for segmenting complex and variable structures. However, as this chapter emphasizes, DL is not always the fastest or the best approach. Classical image processing techniques or pixel classifiers often provide faster, simpler, and highly effective alternatives in many scenarios.

The decision to use DL should be driven by the complexity of the task, the availability of annotated data, and the specific goals of the segmentation project. Successful DL implementations often require significant investments in data curation, annotation, and computational resources. Furthermore, training from scratch is frequently avoidable thanks to the growing ecosystem of pre-trained models and resources shared by the community.

Notably, the landscape of DL segmentation is rapidly evolving. The emergence of foundation models, which are large, versatile networks pre-trained on vast and diverse datasets, promises to further lower the barriers to entry [3]. These models enable transfer learning, fine-tuning, and even zero-shot segmentation, where accurate predictions can be made on previously unseen

data with minimal or no task-specific training. This shift opens exciting new avenues for researchers who previously lacked the resources or expertise to apply DL in their work.

The ongoing development and democratization of DL tools, along with enhancements in model generalisability, human-in-the-loop workflows, and reproducibility, are changing how microscopy data is analyzed. Still, the key to successful segmentation will always involve careful planning, quality control, and selecting the right tool for the task, whether it involves DL or not.

# 10 Output Quality

How to pick a model?

Starting prompt for this chapter: Chapter 10 addresses how to assess the quality of a model’s output, mentioning Metrics Reloaded. This chapter should address the question: how do I know my model is good enough? It should frame this discussion using the example of a segmentation model and discuss how tools can identify uncertain decisions from a model.

- why do we need metrics? to pick models from different training iterations (avoid overfitting) and to select models with different hyperparameters
- mention train/validate/test, discuss the need for “ground truth” data, can callback to chapter 5

## 10.1 Metrics and Losses

- difference between “metric” and “loss”
- loss:
  - has to be “differentiable”
  - used to train the network
  - should already be close to the metric you want to use
  - can be used to assess model quality on validation data (but a custom metric might be more insightful)
- metric:
  - an application specific measure of how close you are to the ground truth
  - used to select a model and to measure progress
  - does not need to be “differentiable”
  - but if it is, that’s great, you can use it as a loss
  - otherwise, find a loss that is a good proxy or chose your model based on the metric on the validation dataset

## 10.2 What Metric to Pick?

- ideally: metric reflects *time/cost needed to clean up* for a *particular application*
- this can mean different things, and sometimes there is no single number (see Cell Tracking below)
- Discuss tradeoffs in metrics choice
- Interpretation of different metrics with guidance on how to balance/trade-off priorities
  - E.g. segmentation for counting (segmentation as object detection) vs. segmentation for size estimation

## 10.3 Examples

Note: We are considering picking one example to place in call out boxes throughout the chapter to facilitate an ongoing discussion with a concrete example. For examples, we will use real data/GT, but synthetically generate the predictions in order to better highlight specific issues.

### 10.3.1 Segmentation Metrics

- show on examples:
  - dice
  - Hausdorff
  - AP\_x (and required matching)
- mention MetricsReloaded

### 10.3.2 Cell Tracking Metrics

- usually not a single number: topological correctness vs. positional correctness
- show traccuracy and discuss some of their metrics

# 11 Outlook

What can AI enable for biology?

Under your first header, include a brief introduction to your chapter.

Starting prompt for this chapter: Chapter 11 concludes the book with a forward-looking assessment of AI in microscopy, focusing on how it can/will enable biological discovery. It should highlight a few motivational examples of discoveries that AI has already enabled and discuss opportunities to which the reader is primed to contribute after reading this book.

## 11.1 Include section headers as appropriate

Use markdown heading level two for section headers. You can use standard markdown formatting, for example *emphasize the end of this sentence*.

This is a new paragraph with more text. Your paragraphs can cross reference other items, such as Figure 11.1. Use `fig` to reference figures, and `eq` to reference equations, such as Equation 11.1.

### 11.1.1 Sub-subsection headers are also available

To make your sections cross reference-able throughout the book, include a section reference, as shown in the header for Section 11.4.

## 11.2 Bibliography and Citations

To cite a research article, add it to references.bib and then refer to the citation key. For example, reference Stringer et al. [107] refers to CellPose and reference Chamier et al. [16] refers to ZeroCostDL4Mic.

## 11.3 Adding to the Glossary

We are using R code to create a glossary for this book. To add a definition, edit the glossary.yml file. To reference the glossary, enclose the word as in these examples: LLMs suffer from [hallucinations](#). It is important to understand the underlying [training data](#), validation data and [false positives](#) to interpret your results. Clicking on the word will reveal its definition by taking you to the entry on the [Glossary](#) page. Pressing back in your browser will return you to your previous place in the textbook.

## 11.4 Code and Equations

This is an example of including a python snippet that generates a figure

```
import matplotlib.pyplot as plt
plt.plot([1,23,2,4])
plt.show()
```

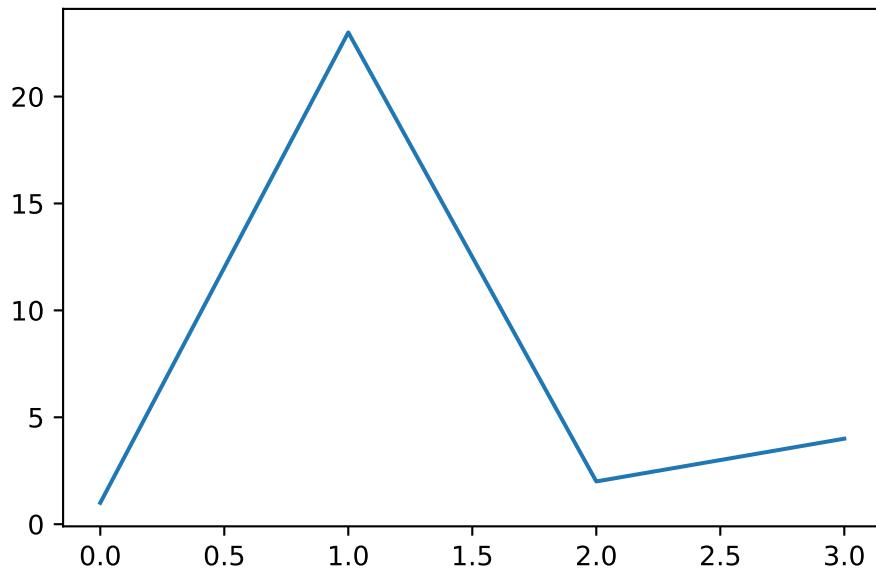


Figure 11.1: Simple Plot

In some cases, you may want to include a code-block that is not executed when the book is compiled. Use the `eval: false` option for this.

```
import matplotlib.pyplot as plt
plt.plot([1,23,2,4])
plt.show()
```

Figures can also be generated that do not show the code by using the option for `code-fold: true`.

```
import numpy as np
import matplotlib.pyplot as plt

r = np.arange(0, 2, 0.01)
theta = 2 * np.pi * r
fig, ax = plt.subplots(
    subplot_kw = {'projection': 'polar'}
)
ax.plot(theta, r)
ax.set_rticks([0.5, 1, 1.5, 2])
ax.grid(True)
plt.show()
```

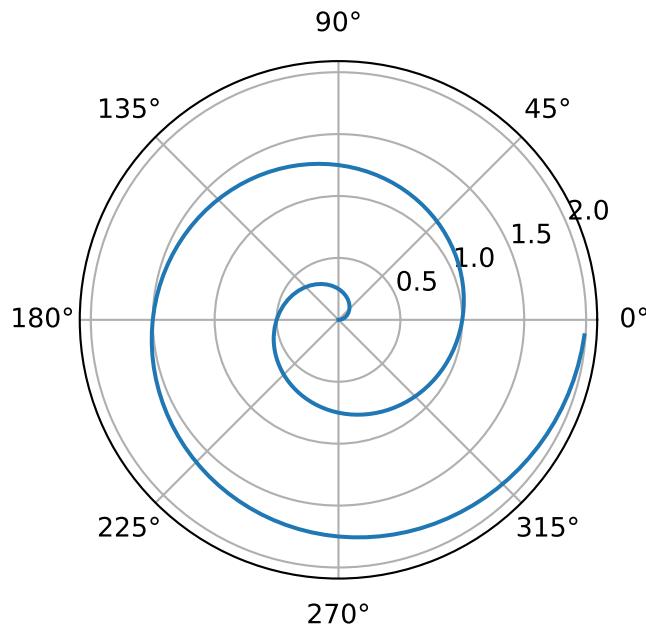


Figure 11.2: A spiral on a polar axis

Here is an example equation.

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (11.1)$$

### 11.4.1 Embedding Figures

You can also embed figures from other notebooks in the repo as shown in the following embed example.

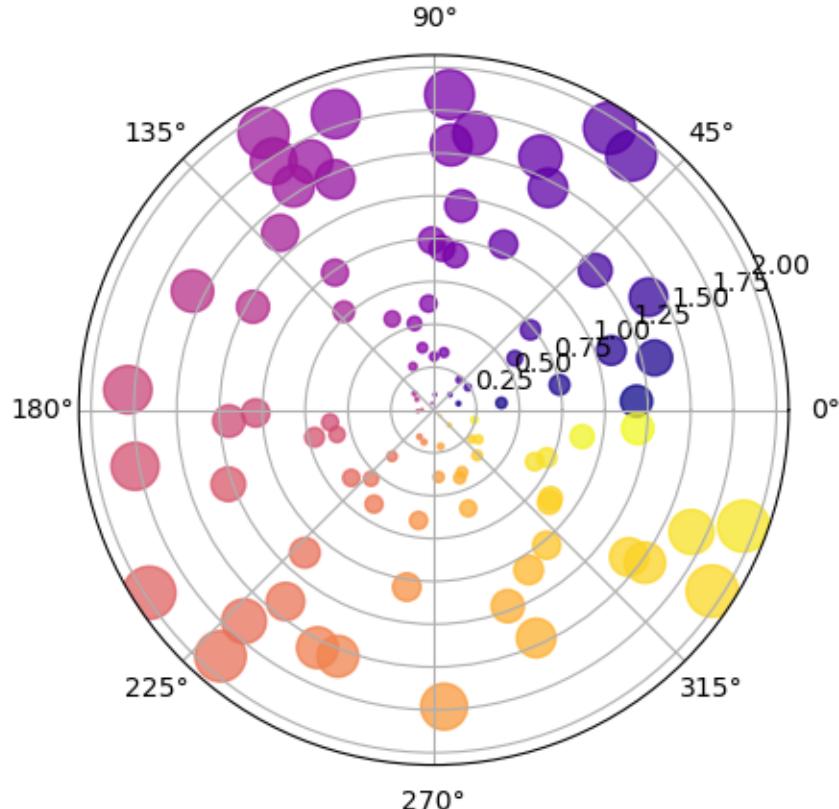


Figure 11.3: Polar plot of circles of random areas at random coords

When embedding notebooks, please store the .ipynb file in the notebook directory. Include the chapter in the name of your file. For example, `chapter4_example_u-net.ipynb`. This is how we will handle chapter- or example-specific environments. We will host notebooks on Google Colab so that any required packages for the code—but not for rendering the book at large—will be installed there. That way, we will not need to handle a global environment across the book.

## 11.5 Quarto has additional features.

You can learn more about markdown options and additional Quarto features in the [Quarto documentation](#). One example that you might find interesting is the option to include callouts in your text. These callouts can be used to highlight potential pitfalls or provide additional optional exercises that the reader might find helpful. Below are examples of the types of callouts available in Quarto.

### Note

Note that there are five types of callouts, including: `note`, `tip`, `warning`, `caution`, and `important`. They can default to open (like this example) or collapsed (example below).

### Tip

These could be good for extra material or exercises.

### Caution

There are caveats when applying these tools. Expand the code below to learn more.

```
r = np.arange(0, 2, 0.01)
theta = 2 * np.pi * r
```

### Warning

Be careful to avoid hallucinations.

### Important

This is key information.

# **Glossary**

## **Aberration**

The spreading of light (also called ‘wavefront distortion’) due to imperfections in the optical path or variations in refractive index at the sample, which results in images that are blurrier than the ideal diffraction-limited image we would expect were aberrations absent.

## **Adaptive Optics**

Technology that senses distortions in the wavefront of light and cancels them, thereby suppressing optical aberrations to enhance image clarity.

## **Autoencoder**

A deep learning architecture used to learn efficient coding of unlabeled data. An autoencoder learns two functions: an encoding function that transforms the input data, and a decoding function that recreates the input data from the encoded representation.

## **Backpropagation**

The method used by neural networks to learn from its predictions. Once the prediction is done, it is compared with the ground truth through a training loss and the value of the comparison is used backwards to sequentially update the weights in the neural network, reward it when making a good prediction and punish it when making a bad prediction.

## **Batch**

A small group of data that is processed together at the same time. For example, when training a machine learning model, a batch is a group of data that is given to the model for learning. Batches are commonly used to make the processes more efficient.

## **Bayesian Optimization**

A strategy that allows the optimization of black-box functions such as deep neural networks. It creates a surrogate model, which is a probabilistic representation of the objective function, using only a few example points.

## **Binary Segmentation**

A type of image segmentation where each pixel is classified into one of two categories—typically “foreground” (e.g., cell) or “background.” The output is a binary mask distinguishing objects (set to a value of 1) from their background (0).

## **CARE (Content-aware image restoration)**

A deep learning-based method for image restoration that leverages content-specific features to enhance degraded images. See <https://github.com/CSBDeep/CSBDeep> for more information.

## **Computer Vision**

A field of computer science wherein computers extract information from images. It often involves object detection within images and can involve classification of the images and/or objects.

## **Convolution**

A mathematical process where a kernel (small matrix) slides over input data (e.g., images) to compute feature maps, highlighting patterns like edges or textures.

## **Convolutional Neural Networks (CNNs)**

A deep learning architecture that applies convolutions to automatically learn features from images for computer vision tasks like classification and detection.

## **Data Augmentation**

A strategy to artificially increase the diversity of a dataset prior to training by applying transformations such as rotation, flipping, or brightness adjustment. It helps improve model robustness and generalisation.

## **Deconvolution**

A mathematical process to partially reverse the blurring effect caused by the microscope's PSF, increasing contrast and resolution over the raw image data if performed carefully.

## **Domain Randomization**

Using simulations or synthetic training data, domain randomization applies random and exaggerated variations to background, lighting, shapes, or textures in the synthetic dataset. This strategy helps the model learn domain-invariant features and is usually used for pretraining a neural network or to enable simulation-to-real transfer.

## **Effect Size**

How "strong" a phenotype is, or how mathematically possible it is to distinguish a given population from the control population.

## **Epoch**

One complete pass through the entire training dataset during the training process.

## **F1 Score**

A classification metric that gives the harmonic mean of precision (proportion of correct true positive predictions across all predicted positive cases) and recall (proportion of true positive predictions against the total positive cases). The harmonic mean is a method to balance both metrics equally. This metric was originally designed for binary classification but can be adapted to multiclass classification by calculating the F1 score per class.

## **False Negatives**

In a scenario where you have two classes “positive” and “negative”, you try to predict cases as one of those classes. False negatives are the cases that you incorrectly predicted as negative and were really positive.

## **False Positives**

In a scenario where you have two classes “positive” and “negative”, you try to predict cases as one of those classes. False positives are the cases that you incorrectly predicted as positive and were really negative.

## **FIJI**

An image processing platform that comes bundled with many plugins for scientific image analysis. See <https://imagej.net/software/fiji/> for more information.

## **Frequency Domain**

The representation of an image as a function of spatial frequency, obtained by transforming an image into the spatial domain using the Fourier transform.

## **Gaussian Process**

A common surrogate model for optimization strategies such as Bayesian Optimization. Gaussian Processes are non-parametric a case that models a conditional probability function. In the hyperparameter search scenario, the Gaussian Process models the probability of getting an objective function value based on some hyperparameters.

## **Generative Adversarial Networks (GANs)**

A deep learning architecture where two neural networks, a generator and a discriminator, are trained in an adversarial process, enabling the generator to create synthetic data, such as realistic images, by learning to deceive the discriminator.

## **Genetic Algorithms**

An optimisation method inspired by the principles of natural selection and genetics. It starts with a population of solutions. These solutions are combined through a process called crossover to produce new solutions (offspring). During this process, random changes or mutations may occur to introduce diversity. After crossover and mutation, a selection step chooses the best solutions from both the parent and offspring populations to form the next generation. This cycle repeats for a set number of generations or until a predefined goal or stopping criterion is met.

## **Ground Truth**

Accurate data against which a model can be evaluated. Ground truth data is often manually annotated. The data type itself will vary depending on the task and evaluation. e.g. instance segmentation may be compared to ground truth object counts or masks.

## **Hallucinations**

Outputs from a model that do not have a basis in the input data and may contain false or misleading information.

## **Hyperparameters**

The options you choose when training a machine learning model that affect the training process or the architecture of the model (e.g., learning rate, batch size, number of layers, training loss, etc.) are called hyperparameters. This term is used to differentiate them from the parameters (also known as weights) of the machine learning model.

## **Image Classification**

A computer vision task where each image is associated with one class and the goal of this task is to correctly predict that class.

## **Image Restoration**

The process of recovering clear, high-quality images from degraded raw data contaminated by blur, noise, or other distortions.

## **Instance Segmentation**

A segmentation task that not only separates objects from the background but also distinguishes between individual objects of the same type (e.g., separating touching cells one by one).

## **IoU**

“Intersection over Union”. A segmentation metric that calculates the difference between the area of overlap between two segmentation masks divided by the area of union.

## **Manual Annotation**

The process of manually labeling specific structures or objects in an image using drawing tools. Typically done in software like Fiji or Napari, this step is essential for creating ground truth data to train or evaluate machine learning models.

## **Metadata**

Any data that provides additional information about other data. In bioimaging, examples include information about sample preparation, the imaging instrument, and image acquisition parameters.

## **N2N (Noise2Noise)**

A supervised denoising method that trains a neural network on pairs of independently noisy images of the same scene, requiring no clean reference data but needing paired noisy inputs. See <https://github.com/NVlabs/noise2noise> for more information.

## **N2S (Noise2Self)**

A self-supervised denoising method that trains a neural network assuming statistically independent noise across the image, requiring only single noisy images without paired clean data. See <https://github.com/czbiohub-sf/noise2self> for more information.

## **N2V (Noise2Void)**

A self-supervised denoising method that trains a neural network to predict pixel values from noisy images by masking input pixels, requiring only single noisy images without paired clean data. See <https://github.com/juglab/n2v> for more information.

## **Nonlinear Problem**

A mathematical problem where the governing equations or operations are nonlinear, meaning outputs are not linearly proportional to inputs.

## **Object Detection**

A computer vision task that identifies and locates individual objects within an image, typically by drawing bounding boxes around them. It provides both the category (what) and position (where) of each object.

## **Panoptic Segmentation**

A computer vision technique that is a combination of semantic segmentation and instance segmentation. It separates an image into regions while also detecting individual object instances within those regions.

## **Pixel Classifiers**

Machine learning models that classify each pixel in an image based on features such as intensity, texture, or local neighborhood. Commonly used in traditional workflows for segmentation or classification tasks.

## **Point Spread Function (PSF)**

A mathematical function that describes how an imaging system blurs a point source.

## **Quality Control Metric**

Any metric that can be used to evaluate quality. It will vary depending on the task and data type. It can be binary (e.g. an image doesn't have debris) or continuous (e.g. annotated object centroids are within 5 pixels of the ground truth centroids).

## **RCAN (residual channel attention network)**

A deep learning-based method using residual learning and channel attention to improve image restoration tasks. See <https://github.com/AiviaCommunity/3D-RCAN> for more information.

## **ReLU**

An activation function common in deep learning that outputs the input directly if it is positive, and outputs zero otherwise; this characteristic helps introduce non-linearity into the model and mitigate the vanishing gradient problem.

## **Self-supervised learning**

A deep learning method where models generate their own supervisory signals from unlabeled data, often by using pretext tasks, to learn useful representations that can be applied to various downstream tasks.

## **Semantic Segmentation**

A form of segmentation where each pixel in an image is assigned to a class (e.g., nucleus, cytoplasm, background), but it does not distinguish between separate instances of the same class.

## **Sigmoid Function**

An activation function common in deep learning that non-linearly maps real inputs to outputs between 0 and 1, being most sensitive to changes in inputs around zero and increasingly compressing extreme positive or negative inputs as they approach 1 or 0 respectively; this characteristic enables it to model probabilities for binary classification and introduce smooth non-linearity.

## **Spatial Domain**

The representation of an image as a function of spatial coordinates.

## **Star-convex Polygon**

A geometric shape used in segmentation algorithms like StarDist. Imagine drawing straight lines (rays) from the centre of an object out toward its edges—if you can see the edge from the centre in all directions, the object is considered star-convex. This method works well for blob-like structures such as nuclei, because their general shape can be captured by measuring how far each ray travels from the centre to the boundary.

## **Supervised learning**

A deep learning method where models learn from labeled data (input-output pairs), enabling them to learn a mapping function for making predictions or decisions on unseen inputs.

## **Training Data**

Data used to train an algorithm to make predictions.

## **Transfer Learning**

A deep learning technique that reuses a model pre-trained on one task as the starting point for a new, related task, leveraging its learned knowledge to improve performance or reduce training requirements. In practice, part of a pretrained neural network (usually the initial layers, responsible for feature extraction) is frozen and reused in a new model. These frozen layers, with the knowledge from a previous dataset, are combined with untrained layers tailored

for a specific bioimaging task. During training, only the new layers will be updated, allowing the model to adapt to the new task with limited data.

## **Transformer Models**

A deep learning architecture based on the multi-head attention mechanism; specifically referring to the ‘vision transformer’ architecture. A vision transformer (ViT) is a transformer designed for computer vision. A ViT decomposes an input image into a series of patches (rather than text into tokens), serializes each patch into a vector, and maps it to a smaller dimension with a single matrix multiplication. These vector embeddings are then processed by a transformer encoder as if they were token embeddings.

## **True Negatives**

In a scenario where you have two classes “positive” and “negative”, you try to predict cases as one of those classes. True positives are the cases that you predicted as negative and were really negative.

## **True Positives**

In a scenario where you have two classes “positive” and “negative”, you try to predict cases as one of those classes. True positives are the cases that you predicted as positive and were really positive.

## **Virtual Machine**

On a physical computer, you install an operating system (e.g., Windows or Ubuntu) that you interact with. A virtual machine is a program that simulates a complete computer with its own operating system. This lets you run a “computer inside your computer” (e.g., using Linux inside Windows or the other way around). As this simulated computer is separate from your physical one, it adds an extra layer of security, because unless the user specifically allows it, the virtual machine cannot access or connect to your real computer.

## **Zernike Modes**

A set of orthogonal polynomials used to describe and correct wavefront aberrations in optical systems.

# References

- [1] Jannis Ahlers et al. *napari: a multi-dimensional image viewer for Python*. July 2023. DOI: [10.5281/zenodo.8115575](https://doi.org/10.5281/zenodo.8115575). URL: <https://zenodo.org/record/8115575> (visited on 10/05/2023).
- [2] Hussain Alibrahim and Simone A. Ludwig. “Hyperparameter Optimization: Comparing Genetic Algorithm against Grid Search and Bayesian Optimization”. In: *2021 IEEE Congress on Evolutionary Computation (CEC)*. 2021, pp. 1551–1559. DOI: [10.1109/CEC45853.2021.9504761](https://doi.org/10.1109/CEC45853.2021.9504761). URL: <https://ieeexplore.ieee.org/document/9504761>.
- [3] Anwai Archit et al. “Segment Anything for Microscopy”. en. In: *Nature Methods* 22.3 (Mar. 2025). Publisher: Nature Publishing Group, pp. 579–591. ISSN: 1548-7105. DOI: [10.1038/s41592-024-02580-4](https://doi.org/10.1038/s41592-024-02580-4). URL: <https://doi.org/10.1038/s41592-024-02580-4> (visited on 04/17/2025).
- [4] Ignacio Arganda-Carreras et al. “Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification”. In: *Bioinformatics* 33.15 (Aug. 2017), pp. 2424–2426. ISSN: 1367-4803. DOI: [10.1093/bioinformatics/btx180](https://doi.org/10.1093/bioinformatics/btx180). URL: <https://doi.org/10.1093/bioinformatics/btx180> (visited on 08/22/2023).
- [5] Matthias Arzt et al. “LABKIT: Labeling and segmentation toolkit for big image data”. In: *Front. Comput. Sci.* 4 (Feb. 2022). ISSN: 2624-9898. URL: <https://www.frontiersin.org/articles/10.3389/fcomp.2022.777728> (visited on 08/30/2022).
- [6] Harikrushnan Balasubramanian et al. “Imagining the future of optical microscopy: everything, everywhere, all at once”. en. In: *Communications Biology* 6.1 (Oct. 2023). Publisher: Nature Publishing Group, p. 1096. ISSN: 2399-3642. DOI: [10.1038/s42003-023-05468-9](https://doi.org/10.1038/s42003-023-05468-9). URL: <https://www.nature.com/articles/s42003-023-05468-9> (visited on 12/03/2025).
- [7] Peter Bankhead et al. “QuPath: Open source software for digital pathology image analysis”. en. In: *Sci. Rep.* 7.1 (Dec. 2017). Number: 1 Publisher: Nature Publishing Group, p. 16878. ISSN: 2045-2322. DOI: [10.1038/s41598-017-17204-5](https://doi.org/10.1038/s41598-017-17204-5). URL: <https://www.nature.com/articles/s41598-017-17204-5> (visited on 10/11/2023).
- [8] Joshua Batson and Loic Royer. “Noise2Self: Blind Denoising by Self-Supervision”. In: *Proceedings of the 36th International Conference on Machine Learning*. Ed. by Kamalika Chaudhuri and Ruslan Salakhutdinov. Vol. 97. Proceedings of Machine Learning Research. PMLR, June 2019, pp. 524–533. URL: <https://proceedings.mlr.press/v97/batson19a.html>.

- [9] Pierre-Louis Bazin et al. “Volumetric neuroimage analysis extensions for the MIPAV software package”. In: *Journal of Neuroscience Methods* 165.1 (2007), pp. 111–121. ISSN: 0165-0270. DOI: <https://doi.org/10.1016/j.jneumeth.2007.05.024>. URL: <https://www.sciencedirect.com/science/article/pii/S0165027007002270>.
- [10] Amir Beck and Marc Teboulle. “A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems”. In: *SIAM Journal on Imaging Sciences* 2.1 (2009), pp. 183–202. DOI: <10.1137/080716542>. eprint: <https://doi.org/10.1137/080716542>. URL: <https://doi.org/10.1137/080716542>.
- [11] Leire Bejarano et al. “Interrogation of endothelial and mural cells in brain metastasis reveals key immune-regulatory mechanisms”. In: *Cancer Cell* 42.3 (Mar. 2024), 378–395.e10. ISSN: 1535-6108. DOI: <10.1016/j.ccr.2023.12.018>. URL: <https://www.sciencedirect.com/science/article/pii/S1535610823004464> (visited on 04/22/2025).
- [12] Stuart Berg et al. “Ilastik: Interactive machine learning for (bio)image analysis”. en. In: *Nat. Methods* 16.12 (Dec. 2019). Number: 12 Publisher: Nature Publishing Group, pp. 1226–1232. ISSN: 1548-7105. DOI: <10.1038/s41592-019-0582-9>. URL: <https://www.nature.com/articles/s41592-019-0582-9> (visited on 08/22/2023).
- [13] Martin J. Booth. “Adaptive Optics in Microscopy”. In: *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* 365.1861 (2007), pp. 2829–2843. ISSN: 1364503X. URL: <http://www.jstor.org/stable/25190627> (visited on 08/19/2025).
- [14] Catherine Bouchard et al. “Resolution enhancement with a task-assisted GAN to guide optical nanoscopy image analysis and acquisition”. In: *Nature Machine Intelligence* 5.8 (Aug. 2023), pp. 830–844. ISSN: 2522-5839. DOI: <10.1038/s42256-023-00689-3>. URL: <https://doi.org/10.1038/s42256-023-00689-3>.
- [15] Juan C. Caicedo et al. “Weakly Supervised Learning of Single-Cell Feature Embeddings”. In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. ISSN: 2575-7075. June 2018, pp. 9309–9318. DOI: <10.1109/CVPR.2018.00970>. URL: <https://ieeexplore.ieee.org/document/8579068> (visited on 04/17/2025).
- [16] Lucas von Chamier et al. “Democratising deep learning for microscopy with Zero-CostDL4Mic”. en. In: *Nature Communications* 12.1 (Apr. 2021), p. 2276. ISSN: 2041-1723. DOI: <10.1038/s41467-021-22518-0>. URL: <https://www.nature.com/articles/s41467-021-22518-0> (visited on 06/24/2022).
- [17] Fabrice de Chaumont et al. “Icy: an open bioimage informatics platform for extended reproducible research”. en. In: *Nat. Methods* 9.7 (June 2012), pp. 690–696. URL: <https://doi.org/10.1038/nmeth.2075>.
- [18] Fei Chen, Paul W. Tillberg, and Edward S. Boyden. “Expansion microscopy”. In: *Science* 347.6221 (2015), pp. 543–548. DOI: <10.1126/science.1260088>. eprint: <https://www.science.org/doi/pdf/10.1126/science.1260088>. URL: <https://www.science.org/doi/abs/10.1126/science.1260088>.

- [19] Huanhuan Chen et al. “Advancements and Practical Considerations for Biophysical Research: Navigating the Challenges and Future of Super-resolution Microscopy”. In: *Chemical & Biomedical Imaging* 2.5 (May 2024), pp. 331–344. DOI: [10.1021/cbmi.4c00019](https://doi.org/10.1021/cbmi.4c00019). URL: <https://doi.org/10.1021/cbmi.4c00019>.
- [20] Jianxu Chen et al. *The Allen Cell and Structure Segmenter: a new open source toolkit for segmenting 3D intracellular structures in fluorescence microscopy images.* en. Pages: 491035 Section: New Results. Dec. 2020. DOI: [10.1101/491035](https://doi.org/10.1101/491035). URL: <https://www.biorxiv.org/content/10.1101/491035v2> (visited on 04/22/2025).
- [21] Jiji Chen et al. “Three-dimensional residual channel attention networks denoise and sharpen fluorescence microscopy image volumes”. In: *Nature Methods* 18.6 (June 2021), pp. 678–687. ISSN: 1548-7105. DOI: [10.1038/s41592-021-01155-x](https://doi.org/10.1038/s41592-021-01155-x). URL: <https://doi.org/10.1038/s41592-021-01155-x>.
- [22] Beth A Cimini. *When to say ‘good enough’*. Oct. 2019. URL: <https://carpenter-singh-lab.broadinstitute.org/blog/when-to-say-good-enough>.
- [23] Ryan Conrad and Kedar Narayan. “Instance segmentation of mitochondria in electron microscopy images with a generalist deep learning model trained on a diverse dataset”. In: *Cell Systems* 14.1 (Jan. 2023), 58–71.e5. ISSN: 2405-4712. DOI: [10.1016/j.cels.2022.12.006](https://doi.org/10.1016/j.cels.2022.12.006). URL: <https://www.sciencedirect.com/science/article/pii/S240547122200494X> (visited on 04/22/2025).
- [24] Kostadin Dabov et al. “Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering”. In: *IEEE Transactions on Image Processing* 16.8 (2007), pp. 2080–2095. DOI: [10.1109/TIP.2007.901238](https://doi.org/10.1109/TIP.2007.901238). URL: <https://ieeexplore.ieee.org/document/4271520>.
- [25] Aram Danielyan et al. “Denoising of two-photon fluorescence images with Block-Matching 3D filtering”. In: *Methods* 68.2 (2014). Tools and Methods for Cellular Localization and Measurements, pp. 308–316. ISSN: 1046-2023. DOI: <https://doi.org/10.1016/j.ymeth.2014.03.010>. URL: <https://www.sciencedirect.com/science/article/pii/S1046202314001030>.
- [26] Elnaz Fazeli et al. “Automated cell tracking using StarDist and TrackMate”. eng. In: *F1000Research* 9 (2020), p. 1279. ISSN: 2046-1402. DOI: [10.12688/f1000research.27019.1](https://doi.org/10.12688/f1000research.27019.1). URL: <https://f1000research.com/articles/9-1279/v1>.
- [27] Francesco Fersini et al. “Wavefront estimation through structured detection in laser scanning microscopy”. In: *Biomed. Opt. Express* 16.5 (May 2025), pp. 2135–2155. DOI: [10.1364/BOE.559899](https://doi.org/10.1364/BOE.559899). URL: <https://opg.optica.org/boe/abstract.cfm?URI=boe-16-5-2135>.
- [28] Daniel Fisch et al. “Molecular definition of the endogenous Toll-like receptor signalling pathways”. en. In: *Nature* 631.8021 (July 2024). Publisher: Nature Publishing Group, pp. 635–644. ISSN: 1476-4687. DOI: [10.1038/s41586-024-07614-7](https://doi.org/10.1038/s41586-024-07614-7). URL: <https://www.nature.com/articles/s41586-024-07614-7> (visited on 04/22/2025).

- [29] Gautier Follain et al. *Fast label-free live imaging reveals key roles of flow dynamics and CD44-HA interaction in cancer cell arrest on endothelial monolayers*. en. Pages: 2024.09.30.615654 Section: New Results. Oct. 2024. doi: [10.1101/2024.09.30.615654](https://doi.org/10.1101/2024.09.30.615654). URL: <https://www.biorxiv.org/content/10.1101/2024.09.30.615654v1> (visited on 03/21/2025).
- [30] Thibaut Goldsborough et al. “InstanSeg: an embedding-based instance segmentation algorithm optimized for accurate, efficient and portable cell segmentation”. In: *arXiv [cs.CV]* (Aug. 2024). URL: <https://doi.org/10.48550/arXiv.2408.15954>.
- [31] Estibaliz Gómez-de-Mariscal et al. “DeepImageJ: A user-friendly environment to run deep learning models in ImageJ”. en. In: *Nat. Methods* 18.10 (Oct. 2021), pp. 1192–1195. URL: <https://doi.org/10.1038/s41592-021-01262-9>.
- [32] R.C. Gonzalez and R.E. Woods. *Digital Image Processing*. Prentice Hall, 2008. ISBN: 9780131687288. URL: <https://books.google.com/books?id=8uGOjRGEzoC>.
- [33] Paul C. Goodwin. “Chapter 10 - Quantitative deconvolution microscopy”. In: *Quantitative Imaging in Cell Biology*. Ed. by Jennifer C. Waters and Torsten Wittman. Vol. 123. Methods in Cell Biology. Academic Press, 2014, pp. 177–192. doi: <https://doi.org/10.1016/B978-0-12-420138-5.00010-0>. URL: <https://www.sciencedirect.com/science/article/pii/B9780124201385000100>.
- [34] Min Guo et al. “Deep learning-based aberration compensation improves contrast and resolution in fluorescence microscopy”. In: *Nature Communications* 16.1 (Jan. 2025), p. 313. ISSN: 2041-1723. doi: [10.1038/s41467-024-55267-x](https://doi.org/10.1038/s41467-024-55267-x). URL: <https://doi.org/10.1038/s41467-024-55267-x>.
- [35] Min Guo et al. “Rapid image deconvolution and multiview fusion for optical microscopy”. In: *Nature Biotechnology* 38.11 (Nov. 2020), pp. 1337–1346. ISSN: 1546-1696. doi: [10.1038/s41587-020-0560-x](https://doi.org/10.1038/s41587-020-0560-x). URL: <https://doi.org/10.1038/s41587-020-0560-x>.
- [36] Robert Haase et al. “A call for FAIR and open-access training materials to advance BioImage Analysis”. In: (Mar. 2024). URL: [https://osf.io/preprints/osf/2zgmc\\_v1](https://osf.io/preprints/osf/2zgmc_v1).
- [37] Guy M Hagen et al. “Fluorescence microscopy datasets for training deep neural networks”. In: *GigaScience* 10.5 (May 2021), giab032. ISSN: 2047-217X. doi: [10.1093/gigascience/giab032](https://doi.org/10.1093/gigascience/giab032). eprint: <https://academic.oup.com/gigascience/article-pdf/10/5/giab032/60688420/giab032.pdf>. URL: <https://doi.org/10.1093/gigascience/giab032>.
- [38] Karen M. Hampson et al. “Adaptive optics for high-resolution imaging”. In: *Nature Reviews Methods Primers* 1.1 (Oct. 2021), p. 68. ISSN: 2662-8449. doi: [10.1038/s43586-021-00066-7](https://doi.org/10.1038/s43586-021-00066-7). URL: <https://doi.org/10.1038/s43586-021-00066-7>.
- [39] Larissa Heinrich et al. “Whole-cell organelle segmentation in volume electron microscopy”. en. In: *Nature* 599.7883 (Nov. 2021). Publisher: Nature Publishing Group, pp. 141–146. ISSN: 1476-4687. doi: [10.1038/s41586-021-03977-3](https://doi.org/10.1038/s41586-021-03977-3). URL: <https://www.nature.com/articles/s41586-021-03977-3> (visited on 04/17/2025).

- [40] Stefan W. Hell. “Far-Field Optical Nanoscopy”. In: *Science* 316.5828 (2007), pp. 1153–1158. DOI: [10.1126/science.1137395](https://doi.org/10.1126/science.1137395). eprint: <https://www.science.org/doi/pdf/10.1126/science.1137395>. URL: <https://www.science.org/doi/abs/10.1126/science.1137395>.
- [41] Iván Hidalgo-Cenalmor et al. “DL4MicEverywhere: deep learning for microscopy made flexible, shareable and reproducible”. en. In: *Nat. Methods* 21.6 (June 2024), pp. 925–927. ISSN: 1548-7091, 1548-7105. DOI: [10.1038/s41592-024-02295-6](https://doi.org/10.1038/s41592-024-02295-6). URL: <https://www.nature.com/articles/s41592-024-02295-6> (visited on 05/20/2024).
- [42] Xuekai Hou et al. “HD2Net: a deep learning framework for simultaneous denoising and deaberration in fluorescence microscopy”. In: *Opt. Express* 33.13 (June 2025), pp. 27317–27333. DOI: [10.1364/OE.554927](https://doi.org/10.1364/OE.554927). URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-33-13-27317>.
- [43] Lejia Hu et al. “Image enhancement for fluorescence microscopy based on deep learning with prior knowledge of aberration”. In: *Opt. Lett.* 46.9 (May 2021), pp. 2055–2058. DOI: [10.1364/OL.418997](https://doi.org/10.1364/OL.418997). URL: <https://opg.optica.org/ol/abstract.cfm?URI=ol-46-9-2055>.
- [44] Ilija Ilievski et al. “Efficient Hyperparameter Optimization for Deep Learning Algorithms Using Deterministic RBF Surrogates”. In: *Proceedings of the AAAI Conference on Artificial Intelligence* 31.1 (Feb. 2017). DOI: [10.1609/aaai.v31i1.10647](https://doi.org/10.1609/aaai.v31i1.10647). URL: <https://ojs.aaai.org/index.php/AAAI/article/view/10647>.
- [45] Nasim Jamali et al. “2020 BioImage Analysis Survey: Community experiences and needs for the future”. In: *Biological Imaging* 1 (Jan. 2021), e4. DOI: [10.1017/S2633903X21000039](https://doi.org/10.1017/S2633903X21000039). URL: <https://www.cambridge.org/core/journals/biological-imaging/article/2020-bioimage-analysis-survey-community-experiences-and-needs-for-the-future/9E824DC0C27568FE5B9D12FB59B1BB90>.
- [46] Na Ji. “Adaptive optical fluorescence microscopy”. In: *Nature Methods* 14.4 (Apr. 2017), pp. 374–380. ISSN: 1548-7105. DOI: [10.1038/nmeth.4218](https://doi.org/10.1038/nmeth.4218). URL: <https://doi.org/10.1038/nmeth.4218>.
- [47] Ziwei Ji, Justin D. Li, and Matus Telgarsky. “Early-stopped neural networks are consistent”. English (US). In: *Advances in Neural Information Processing Systems 34 - 35th Conference on Neural Information Processing Systems, NeurIPS 2021*. Ed. by Marc’Aurelio Ranzato et al. Advances in Neural Information Processing Systems. Publisher Copyright: © 2021 Neural information processing systems foundation. All rights reserved.; 35th Conference on Neural Information Processing Systems, NeurIPS 2021 ; Conference date: 06-12-2021 Through 14-12-2021. Neural information processing systems foundation, 2021, pp. 1805–1817. URL: <https://dl.acm.org/doi/10.5555/3540261.3540400>.
- [48] Iksung Kang et al. “Adaptive optical correction in in vivo two-photon fluorescence microscopy with neural fields”. In: *bioRxiv* (2024). DOI: [10.1101/2024.10.20.619284](https://doi.org/10.1101/2024.10.20.619284). eprint: <https://www.biorxiv.org/content/early/2024/10/22/2024.10.20.619284.full.pdf>. URL: <https://www.biorxiv.org/content/early/2024/10/22/2024.10.20.619284>.

- [49] Iksung Kang et al. “Coordinate-based neural representations for computational adaptive optics in widefield microscopy”. In: *Nature Machine Intelligence* 6.6 (June 2024), pp. 714–725. ISSN: 2522-5839. DOI: [10.1038/s42256-024-00853-3](https://doi.org/10.1038/s42256-024-00853-3). URL: <http://dx.doi.org/10.1038/s42256-024-00853-3>.
- [50] Thomas Kluyver et al. “Jupyter Notebooks - a publishing format for reproducible computational workflows”. In: *Positioning and Power in Academic Publishing: Players, Agents and Agendas*. Ed. by Fernando Loizides and Birgit Schmidt. Netherlands: IOS Press, 2016, pp. 87–90. URL: <https://eprints.soton.ac.uk/403913/>.
- [51] Bogdan Kochetov et al. “UNSEG: unsupervised segmentation of cells and their nuclei in complex tissue samples”. en. In: *Communications Biology* 7.1 (Aug. 2024). Publisher: Nature Publishing Group, pp. 1–14. ISSN: 2399-3642. DOI: [10.1038/s42003-024-06714-4](https://doi.org/10.1038/s42003-024-06714-4). URL: <https://www.nature.com/articles/s42003-024-06714-4> (visited on 04/22/2025).
- [52] Anna Kreshuk and Chong Zhang. “Machine Learning: Advanced Image Segmentation Using ilastik”. en. In: *Methods Mol. Biol.* 2040 (2019), pp. 449–463. URL: [https://link.springer.com/protocol/10.1007/978-1-4939-9686-5\\_21](https://link.springer.com/protocol/10.1007/978-1-4939-9686-5_21).
- [53] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. “ImageNet Classification with Deep Convolutional Neural Networks”. In: *Advances in Neural Information Processing Systems*. Vol. 25. Curran Associates, Inc., 2012. URL: [https://proceedings.neurips.cc/paper\\_files/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html](https://proceedings.neurips.cc/paper_files/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html) (visited on 10/05/2023).
- [54] Alexander Krull, Tim-Oliver Buchholz, and Florian Jug. “Noise2Void - Learning Denoising From Single Noisy Images”. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2019, pp. 2124–2132. DOI: [10.1109/CVPR.2019.00223](https://doi.org/10.1109/CVPR.2019.00223). URL: <https://ieeexplore.ieee.org/document/8954066>.
- [55] Romain F. Laine et al. “Avoiding a replication crisis in deep-learning-based bioimage analysis”. In: *Nature methods* 18.10 (Oct. 2021), pp. 1136–1144. ISSN: 1548-7091. DOI: [10.1038/s41592-021-01284-3](https://doi.org/10.1038/s41592-021-01284-3). URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7611896/> (visited on 10/06/2023).
- [56] Y. Lecun et al. “Gradient-based learning applied to document recognition”. In: *Proceedings of the IEEE* 86.11 (Nov. 1998), pp. 2278–2324. ISSN: 1558-2256. DOI: [10.1109/5.726791](https://doi.org/10.1109/5.726791). URL: <https://ieeexplore.ieee.org/document/726791> (visited on 04/22/2025).
- [57] Rachel M. Lee et al. “Believing is seeing – the deceptive influence of bias in quantitative microscopy”. In: *Journal of Cell Science* 137.1 (Jan. 2024), jcs261567. ISSN: 0021-9533. DOI: [10.1242/jcs.261567](https://doi.org/10.1242/jcs.261567). URL: <https://doi.org/10.1242/jcs.261567> (visited on 12/03/2025).
- [58] Jaakkko Lehtinen et al. “Noise2Noise: Learning Image Restoration without Clean Data”. In: *Proceedings of the 35th International Conference on Machine Learning*. Ed. by Jennifer Dy and Andreas Krause. Vol. 80. Proceedings of Machine Learning Research. PMLR, July 2018, pp. 2965–2974. URL: <https://proceedings.mlr.press/v80/lehtinen18a.html>.

- [59] Jizhou Li, Florian Luisier, and Thierry Blu. “Pure-let deconvolution of 3D fluorescence microscopy images”. In: *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. 2017, pp. 723–727. DOI: [10.1109/ISBI.2017.7950621](https://doi.org/10.1109/ISBI.2017.7950621). URL: <https://ieeexplore.ieee.org/document/7950621>.
- [60] Yue Li et al. “Incorporating the image formation process into deep learning improves network performance”. In: *Nature Methods* 19.11 (Nov. 2022), pp. 1427–1437. ISSN: 1548-7105. DOI: [10.1038/s41592-022-01652-7](https://doi.org/10.1038/s41592-022-01652-7). URL: <https://doi.org/10.1038/s41592-022-01652-7>.
- [61] Yuxiang Li and Linlin Shen. “cC-GAN: A Robust Transfer-Learning Framework for HEp-2 Specimen Image Segmentation”. In: *IEEE Access* 6 (2018), pp. 14048–14058. ISSN: 2169-3536. DOI: [10.1109/ACCESS.2018.2808938](https://doi.org/10.1109/ACCESS.2018.2808938). URL: <https://ieeexplore.ieee.org/document/8301400> (visited on 04/17/2025).
- [62] Binbin Lin et al. “A deep learned nanowire segmentation model using synthetic data augmentation”. en. In: *npj Computational Materials* 8.1 (Apr. 2022). Publisher: Nature Publishing Group, pp. 1–12. ISSN: 2057-3960. DOI: [10.1038/s41524-022-00767-x](https://doi.org/10.1038/s41524-022-00767-x). URL: <https://www.nature.com/articles/s41524-022-00767-x> (visited on 04/22/2025).
- [63] Bojing Liu et al. “Self-supervised learning reveals clinically relevant histomorphological patterns for therapeutic strategies in colon cancer”. en. In: *Nature Communications* 16.1 (Mar. 2025). Publisher: Nature Publishing Group, p. 2328. ISSN: 2041-1723. DOI: [10.1038/s41467-025-57541-y](https://doi.org/10.1038/s41467-025-57541-y). URL: <https://www.nature.com/articles/s41467-025-57541-y> (visited on 04/17/2025).
- [64] Jihong Liu et al. “A Saturation Artifacts Inpainting Method Based on Two-Stage GAN for Fluorescence Microscope Images”. In: *Micromachines* 15.7 (2024). ISSN: 2072-666X. DOI: [10.3390/mi15070928](https://doi.org/10.3390/mi15070928). URL: <https://www.mdpi.com/2072-666X/15/7/928>.
- [65] Vebjorn Ljosa, Katherine L Sokolnicki, and Anne E Carpenter. “Annotated high-throughput microscopy image sets for validation”. en. In: *Nat. Methods* 9.7 (June 2012), p. 637. URL: <https://doi.org/10.1038/nmeth.2083>.
- [66] L. B. Lucy. “An iterative technique for the rectification of observed distributions”. In: *Astronomical Journal* 79 (June 1974), p. 745. DOI: [10.1086/111605](https://doi.org/10.1086/111605).
- [67] Florian Luisier et al. “Fast interscale wavelet denoising of Poisson-corrupted images”. In: *Signal Processing* 90.2 (2010), pp. 415–427. ISSN: 0165-1684. DOI: <https://doi.org/10.1016/j.sigpro.2009.07.009>. URL: <https://www.sciencedirect.com/science/article/pii/S0165168409003016>.
- [68] Markku Makitalo and Alessandro Foi. “Optimal Inversion of the Generalized Anscombe Transformation for Poisson-Gaussian Noise”. In: *IEEE Transactions on Image Processing* 22.1 (2013), pp. 91–103. DOI: [10.1109/TIP.2012.2202675](https://doi.org/10.1109/TIP.2012.2202675). URL: <https://ieeexplore.ieee.org/document/6212354>.
- [69] Keith Miller. “Least Squares Methods for Ill-Posed Problems with a Prescribed Bound”. In: *SIAM Journal on Mathematical Analysis* 1.1 (1970), pp. 52–74. DOI: [10.1137/0501006](https://doi.org/10.1137/0501006). eprint: <https://doi.org/10.1137/0501006>. URL: <https://doi.org/10.1137/0501006>.

- [70] Tamás Miseta, Attila Fodor, and Ágnes Vathy-Fogarassy. “Surpassing early stopping: A novel correlation-based stopping criterion for neural networks”. In: *Neurocomputing* 567 (2024), p. 127028. ISSN: 0925-2312. DOI: <https://doi.org/10.1016/j.neucom.2023.127028>. URL: <https://www.sciencedirect.com/science/article/pii/S0925231223011517>.
- [71] Erick Moen et al. “Deep learning for cellular image analysis”. en. In: *Nature Methods* 16.12 (Dec. 2019). Number: 12 Publisher: Nature Publishing Group, pp. 1233–1246. ISSN: 1548-7105. DOI: <10.1038/s41592-019-0403-1>. URL: <https://www.nature.com/articles/s41592-019-0403-1> (visited on 11/30/2023).
- [72] Mohammad Amin Morid, Alireza Borjali, and Guilherme Del Fiol. “A scoping review of transfer learning research on medical image analysis using ImageNet”. In: *Computers in Biology and Medicine* 128 (Jan. 2021), p. 104115. ISSN: 0010-4825. DOI: <10.1016/j.compbiomed.2020.104115>. URL: <https://www.sciencedirect.com/science/article/pii/S0010482520304467> (visited on 04/17/2025).
- [73] Nikita Moshkov et al. “Learning representations for image-based profiling of perturbations”. en. In: *Nature Communications* 15.1 (Feb. 2024). Publisher: Nature Publishing Group, p. 1594. ISSN: 2041-1723. DOI: <10.1038/s41467-024-45999-1>. URL: <https://www.nature.com/articles/s41467-024-45999-1> (visited on 04/17/2025).
- [74] *napari: a multi-dimensional image viewer for Python*. DOI: <10.5281/zenodo.3555620>. URL: <https://zenodo.org/records/16883660>.
- [75] Kefu Ning et al. “Deep self-learning enables fast, high-fidelity isotropic resolution restoration for volumetric fluorescence microscopy”. In: *Light: Science & Applications* 12.1 (Aug. 2023), p. 204. ISSN: 2047-7538. DOI: <10.1038/s41377-023-01230-2>. URL: <https://doi.org/10.1038/s41377-023-01230-2>.
- [76] Pamela Osuna-Vargas et al. “Denoising Diffusion Models for High-Resolution Microscopy Image Restoration”. In: *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. 2025, pp. 4320–4330. DOI: <10.1109/WACV61041.2025.00424>. URL: <https://ieeexplore.ieee.org/document/10943369>.
- [77] Wei Ouyang et al. “BioImage Model Zoo: A Community-Driven Resource for Accessible Deep Learning in BioImage Analysis”. en. In: *bioRxiv* (June 2022). Pages: 2022.06.07.495102 Section: New Results, p. 2022.06.07.495102. DOI: <10.1101/2022.06.07.495102>. URL: <https://www.biorxiv.org/content/10.1101/2022.06.07.495102v1> (visited on 10/05/2023).
- [78] Wei Ouyang et al. “ImJoy: an open-source computational platform for the deep learning era”. en. In: *Nat. Methods* 16.12 (Dec. 2019), pp. 1199–1200. URL: <https://doi.org/10.1038/s41592-019-0627-0>.
- [79] Marius Pachitariu and Carsen Stringer. “Cellpose 2.0: how to train your own model”. en. In: *Nat. Methods* 19.12 (Dec. 2022). Publisher: Nature Publishing Group, pp. 1634–1641. ISSN: 1548-7105. DOI: <10.1038/s41592-022-01663-4>. URL: <http://www.nature.com/articles/s41592-022-01663-4> (visited on 11/09/2022).

- [80] Eunwoo Park et al. “Unsupervised inter-domain transformation for virtually stained high-resolution mid-infrared photoacoustic microscopy using explainable deep learning”. In: *Nature Communications* 15.1 (Dec. 2024), p. 10892. ISSN: 2041-1723. DOI: [10.1038/s41467-024-55262-2](https://doi.org/10.1038/s41467-024-55262-2). URL: <https://doi.org/10.1038/s41467-024-55262-2>.
- [81] Hyoungjun Park et al. “Deep learning enables reference-free isotropic super-resolution for volumetric fluorescence microscopy”. In: *Nature Communications* 13.1 (June 2022), p. 3297. ISSN: 2041-1723. DOI: [10.1038/s41467-022-30949-6](https://doi.org/10.1038/s41467-022-30949-6). URL: <https://doi.org/10.1038/s41467-022-30949-6>.
- [82] Joanna W. Pylvänäinen et al. “Live-cell imaging in the deep learning era”. en. In: *Current Opinion in Cell Biology* 85 (Dec. 2023), p. 102271. ISSN: 09550674. DOI: [10.1016/j.ceb.2023.102271](https://doi.org/10.1016/j.ceb.2023.102271). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0955067423001205> (visited on 11/07/2023).
- [83] Chang Qiao et al. “Deep learning-based optical aberration estimation enables offline digital adaptive optics and super-resolution imaging”. In: *Photon. Res.* 12.3 (Mar. 2024), pp. 474–484. DOI: [10.1364/PRJ.506778](https://doi.org/10.1364/PRJ.506778). URL: <https://opg.optica.org/prj/abstract.cfm?URI=prj-12-3-474>.
- [84] Chang Qiao et al. “Evaluation and development of deep neural networks for image super-resolution in optical microscopy”. In: *Nature Methods* 18.2 (Feb. 2021), pp. 194–202. ISSN: 1548-7105. DOI: [10.1038/s41592-020-01048-5](https://doi.org/10.1038/s41592-020-01048-5). URL: <https://doi.org/10.1038/s41592-020-01048-5>.
- [85] Chang Qiao et al. “Rationalized deep learning super-resolution microscopy for sustained live imaging of rapid subcellular processes”. In: *Nature Biotechnology* 41.3 (Mar. 2023), pp. 367–377. ISSN: 1546-1696. DOI: [10.1038/s41587-022-01471-3](https://doi.org/10.1038/s41587-022-01471-3). URL: <https://doi.org/10.1038/s41587-022-01471-3>.
- [86] Luis Rangel DaCosta et al. “A robust synthetic data generation framework for machine learning in high-resolution transmission electron microscopy (HRTEM)”. en. In: *npj Computational Materials* 10.1 (July 2024). Publisher: Nature Publishing Group, pp. 1–11. ISSN: 2057-3960. DOI: [10.1038/s41524-024-01336-0](https://doi.org/10.1038/s41524-024-01336-0). URL: <https://www.nature.com/articles/s41524-024-01336-0> (visited on 04/22/2025).
- [87] William Hadley Richardson. “Bayesian-Based Iterative Method of Image Restoration\*”. In: *J. Opt. Soc. Am.* 62.1 (Jan. 1972), pp. 55–59. DOI: [10.1364/JOSA.62.000055](https://doi.org/10.1364/JOSA.62.000055). URL: <https://opg.optica.org/abstract.cfm?URI=josa-62-1-55>.
- [88] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-Net: Convolutional Networks for Biomedical Image Segmentation”. en. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by Nassir Navab et al. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015, pp. 234–241. ISBN: 978-3-319-24574-4. DOI: [10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28). URL: [https://link.springer.com/chapter/10.1007/978-3-319-24574-4\\_28](https://link.springer.com/chapter/10.1007/978-3-319-24574-4_28).

- [89] Curtis T Rueden et al. “Scientific Community Image Forum: A discussion forum for scientific image software”. en. In: *PLoS Biol.* 17.6 (June 2019), e3000340. URL: <https://doi.org/10.1371/journal.pbio.3000340>.
- [90] Craig T. Russell et al. *bia-binder: A web-native cloud compute service for the bioimage analysis community*. arXiv:2411.12662 [q-bio]. Nov. 2024. DOI: [10.48550/arXiv.2411.12662](https://doi.org/10.48550/arXiv.2411.12662). URL: <http://arxiv.org/abs/2411.12662> (visited on 04/09/2025).
- [91] Daniel Sage et al. “DeconvolutionLab2: An open-source software for deconvolution microscopy”. In: *Methods* 115 (2017). Image Processing for Biologists, pp. 28–41. ISSN: 1046-2023. DOI: <https://doi.org/10.1016/j.ymeth.2016.12.015>. URL: <https://www.sciencedirect.com/science/article/pii/S1046202316305096>.
- [92] Debayan Saha et al. “Practical sensorless aberration estimation for 3D microscopy with deep learning”. In: *Opt. Express* 28.20 (Sept. 2020), pp. 29044–29053. DOI: [10.1364/OE.401933](https://doi.org/10.1364/OE.401933). URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-28-20-29044>.
- [93] Steffen J. Sahl, Stefan W. Hell, and Stefan Jakobs. “Fluorescence nanoscopy in cell biology”. In: *Nature Reviews Molecular Cell Biology* 18.11 (Nov. 2017), pp. 685–701. ISSN: 1471-0080. DOI: [10.1038/nrm.2017.71](https://doi.org/10.1038/nrm.2017.71). URL: <https://doi.org/10.1038/nrm.2017.71>.
- [94] Claudio Filipi Gonçalves Dos Santos and João Paulo Papa. “Avoiding Overfitting: A Survey on Regularization Methods for Convolutional Neural Networks”. In: *ACM Comput. Surv.* 54.10s (Sept. 2022). ISSN: 0360-0300. DOI: [10.1145/3510413](https://doi.org/10.1145/3510413). URL: <https://doi.org/10.1145/3510413>.
- [95] P. Sarder and A. Nehorai. “Deconvolution methods for 3-D fluorescence microscopy images”. In: *IEEE Signal Processing Magazine* 23.3 (2006), pp. 32–45. DOI: [10.1109/MSP.2006.1628876](https://doi.org/10.1109/MSP.2006.1628876). URL: <https://ieeexplore.ieee.org/document/1628876>.
- [96] Lothar Schermelleh, Rainer Heintzmann, and Heinrich Leonhardt. “A guide to super-resolution fluorescence microscopy”. In: *Journal of Cell Biology* 190.2 (July 2010), pp. 165–175. ISSN: 0021-9525. DOI: [10.1083/jcb.201002018](https://doi.org/10.1083/jcb.201002018). eprint: [https://rupress.org/jcb/article-pdf/190/2/165/1568325/jcb\\_\\\_201002018.pdf](https://rupress.org/jcb/article-pdf/190/2/165/1568325/jcb_\_201002018.pdf). URL: <https://doi.org/10.1083/jcb.201002018>.
- [97] Lothar Schermelleh et al. “Super-resolution microscopy demystified”. In: *Nature Cell Biology* 21.1 (Jan. 2019), pp. 72–84. ISSN: 1476-4679. DOI: [10.1038/s41556-018-0251-8](https://doi.org/10.1038/s41556-018-0251-8). URL: <https://doi.org/10.1038/s41556-018-0251-8>.
- [98] Johannes Schindelin et al. “Fiji: an open-source platform for biological-image analysis”. en. In: *Nature Methods* 9.7 (July 2012). ISBN: 1548-7105, pp. 676–682. ISSN: 1548-7105. DOI: [10.1038/nmeth.2019](https://doi.org/10.1038/nmeth.2019). URL: <https://doi.org/10.1038/nmeth.2019>.
- [99] Johannes Schindelin et al. “The ImageJ ecosystem: An open platform for biomedical image analysis”. en. In: *Mol. Reprod. Dev.* 82.7-8 (July 2015), pp. 518–529. URL: <https://doi.org/10.1002/mrd.22489>.

- [100] Uwe Schmidt et al. “Cell Detection with Star-Convex Polygons”. en. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*. Ed. by Alejandro F. Frangi et al. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2018, pp. 265–273. ISBN: 978-3-030-00934-2. DOI: [10.1007/978-3-030-00934-2\\_30](https://doi.org/10.1007/978-3-030-00934-2_30). URL: [https://link.springer.com/chapter/10.1007/978-3-030-00934-2\\_30](https://link.springer.com/chapter/10.1007/978-3-030-00934-2_30).
- [101] Rajavi Shah, Nodar Gogoberidze, and Beth Cimini. *Bilayers*. URL: <https://github.com/bilayer-containers/bilayers>.
- [102] Zafran Hussain Shah et al. “Image restoration in frequency space using complex-valued CNNs”. In: *Frontiers in Artificial Intelligence* Volume 7 - 2024 (2024). ISSN: 2624-8212. DOI: [10.3389/frai.2024.1353873](https://doi.org/10.3389/frai.2024.1353873). URL: <https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2024.1353873>.
- [103] Connor Shorten and Taghi M. Khoshgoftaar. “A survey on Image Data Augmentation for Deep Learning”. In: *Journal of Big Data* 6.1 (July 2019), p. 60. ISSN: 2196-1115. DOI: [10.1186/s40537-019-0197-0](https://doi.org/10.1186/s40537-019-0197-0). URL: <https://doi.org/10.1186/s40537-019-0197-0> (visited on 04/22/2025).
- [104] Hari Shroff et al. “Live-cell imaging powered by computation”. In: *Nature Reviews Molecular Cell Biology* 25.6 (June 2024), pp. 443–463. ISSN: 1471-0080. DOI: [10.1038/s41580-024-00702-6](https://doi.org/10.1038/s41580-024-00702-6). URL: <https://doi.org/10.1038/s41580-024-00702-6>.
- [105] Suganya Sivagurunathan et al. “Bridging imaging users to imaging analysis - A community survey”. en. In: *J. Microsc.* (Sept. 2023). URL: <https://doi.org/10.1111/jmi.13229>.
- [106] David R Stirling et al. “CellProfiler 4: improvements in speed, utility and usability”. en. In: *BMC Bioinformatics* 22.1 (Sept. 2021), p. 433. URL: <https://doi.org/10.1186/s12859-021-04344-9>.
- [107] Carsen Stringer et al. “Cellpose: a generalist algorithm for cellular segmentation”. en. In: *Nature Methods* 18.1 (Jan. 2021), pp. 100–106. ISSN: 1548-7105. DOI: [10.1038/s41592-020-01018-x](https://doi.org/10.1038/s41592-020-01018-x). URL: <https://www.nature.com/articles/s41592-020-01018-x> (visited on 06/03/2024).
- [108] A. N. Tikhonov. “Solution of incorrectly formulated problems and the regularization method”. In: *Soviet Math. Dokl.* 4 (1963), pp. 1035–1038.
- [109] Jessica Valli et al. “Seeing beyond the limit: A guide to choosing the right super-resolution microscopy technique”. In: *Journal of Biological Chemistry* 297.1 (2021), p. 100791. ISSN: 0021-9258. DOI: <https://doi.org/10.1016/j.jbc.2021.100791>. URL: <https://www.sciencedirect.com/science/article/pii/S0021925821005846>.
- [110] Manasij Venkatesh, Kavya Mohan, and Chandra Sekhar Seelamantula. “Directional bilateral filters for smoothing fluorescence microscopy images”. In: *AIP Advances* 5.8 (Aug. 2015), p. 084805. ISSN: 2158-3226. DOI: [10.1063/1.4930029](https://doi.org/10.1063/1.4930029). eprint: [https://pubs.aip.org/aip/adv/article-pdf/doi/10.1063/1.4930029/12895317/084805\\\_\\\_1\\\_\\\_online.pdf](https://pubs.aip.org/aip/adv/article-pdf/doi/10.1063/1.4930029/12895317/084805\_\_1\_\_online.pdf). URL: <https://doi.org/10.1063/1.4930029>.

- [111] Giuseppe Vicedomini, Paolo Bianchini, and Alberto Diaspro. “STED super-resolved microscopy”. In: *Nature Methods* 15.3 (Mar. 2018), pp. 173–182. ISSN: 1548-7105. DOI: [10.1038/nmeth.4593](https://doi.org/10.1038/nmeth.4593). URL: <https://doi.org/10.1038/nmeth.4593>.
- [112] Eric C. Wait, Michael A. Reiche, and Teng-Leong Chew. “Hypothesis-driven quantitative fluorescence microscopy – the importance of reverse-thinking in experimental design”. In: *Journal of Cell Science* 133.21 (Nov. 2020), jcs250027. ISSN: 0021-9533. DOI: [10.1242/jcs.250027](https://doi.org/10.1242/jcs.250027). URL: <https://doi.org/10.1242/jcs.250027> (visited on 12/03/2025).
- [113] Stéfan van der Walt et al. “scikit-image: image processing in Python”. en. In: *PeerJ* 2 (June 2014), e453. URL: <https://peerj.com/articles/453/>.
- [114] Hongda Wang et al. “Deep learning enables cross-modality super-resolution in fluorescence microscopy”. In: *Nature Methods* 16.1 (Jan. 2019), pp. 103–110. ISSN: 1548-7105. DOI: [10.1038/s41592-018-0239-0](https://doi.org/10.1038/s41592-018-0239-0). URL: <https://doi.org/10.1038/s41592-018-0239-0>.
- [115] Zhou Wang et al. “Image quality assessment: from error visibility to structural similarity”. In: *IEEE Transactions on Image Processing* 13.4 (2004), pp. 600–612. DOI: [10.1109/TIP.2003.819861](https://doi.org/10.1109/TIP.2003.819861). URL: <https://ieeexplore.ieee.org/document/1284395>.
- [116] Asmamaw T. Wassie, Yongxin Zhao, and Edward S. Boyden. *Expansion microscopy: principles and uses in biological research*. Jan. 2019. DOI: [10.1038/s41592-018-0219-4](https://doi.org/10.1038/s41592-018-0219-4). URL: <https://doi.org/10.1038/s41592-018-0219-4>.
- [117] Martin Weigert et al. “Content-aware image restoration: pushing the limits of fluorescence microscopy”. In: *Nature Methods* 15.12 (Dec. 2018), pp. 1090–1097. ISSN: 1548-7105. DOI: [10.1038/s41592-018-0216-7](https://doi.org/10.1038/s41592-018-0216-7). URL: <https://doi.org/10.1038/s41592-018-0216-7>.
- [118] Norbert Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series: With Engineering Applications*. The MIT Press, Aug. 1949. ISBN: 9780262257190. DOI: [10.7551/mitpress/2946.001.0001](https://doi.org/10.7551/mitpress/2946.001.0001). eprint: [https://direct.mit.edu/book-pdf/2313079/book\\_9780262257190.pdf](https://direct.mit.edu/book-pdf/2313079/book_9780262257190.pdf). URL: <https://doi.org/10.7551/mitpress/2946.001.0001>.
- [119] Yicong Wu and Hari Shroff. “Faster, sharper, and deeper: structured illumination microscopy for biological imaging”. In: *Nature Methods* 15.12 (Dec. 2018), pp. 1011–1019. ISSN: 1548-7105. DOI: [10.1038/s41592-018-0211-z](https://doi.org/10.1038/s41592-018-0211-z). URL: <https://doi.org/10.1038/s41592-018-0211-z>.
- [120] Yicong Wu and Hari Shroff. “Multiscale fluorescence imaging of living samples”. In: *Histochemistry and Cell Biology* 158.4 (Oct. 2022), pp. 301–323. ISSN: 1432-119X. DOI: [10.1007/s00418-022-02147-4](https://doi.org/10.1007/s00418-022-02147-4). URL: <https://doi.org/10.1007/s00418-022-02147-4>.
- [121] Kyrollos Yanny et al. “Deep learning for fast spatially varying deconvolution”. In: *Optica* 9.1 (Jan. 2022), pp. 96–99. DOI: [10.1364/OPTICA.442438](https://doi.org/10.1364/OPTICA.442438). URL: <https://opg.optica.org/optica/abstract.cfm?URI=optica-9-1-96>.
- [122] Chong Zhang et al. “Bio-Image Informatics Index BIII: A unique database of image analysis tools and workflows for and by the bioimaging community”. In: *arXiv [q-bio.QM]* (Dec. 2023). URL: <https://doi.org/10.48550/arXiv.2312.11256>.

- [123] Kai Zhang et al. “Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising”. In: *IEEE Transactions on Image Processing* 26.7 (2017), pp. 3142–3155. DOI: [10.1109/TIP.2017.2662206](https://doi.org/10.1109/TIP.2017.2662206). URL: <https://ieeexplore.ieee.org/document/7839189>.
- [124] Yide Zhang et al. “A Poisson-Gaussian Denoising Dataset With Real Fluorescence Microscopy Images”. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Optica Publishing Group, Sept. 2019, pp. 11702–11710. DOI: [10.1109/CVPR.2019.01198](https://doi.org/10.1109/CVPR.2019.01198). URL: <https://ieeexplore.ieee.org/document/8953965>.
- [125] Yulun Zhang et al. *Image Super-Resolution Using Very Deep Residual Channel Attention Networks*. 2018. arXiv: [1807.02758 \[cs.CV\]](https://arxiv.org/abs/1807.02758). URL: <https://arxiv.org/abs/1807.02758>.
- [126] Liqun Zhong, Guole Liu, and Ge Yang. “Blind Denoising of Fluorescence Microscopy Images Using GAN-Based Global Noise Modeling”. In: *ISBI*. 2021, pp. 863–867. URL: <https://doi.org/10.1109/ISBI48211.2021.9434150>.
- [127] You Zhou et al. “Aberration Modeling in Deep Learning for Volumetric Reconstruction of Light-Field Microscopy”. In: *Laser & Photonics Reviews* 17.10 (2023), p. 2300154. DOI: <https://doi.org/10.1002/lpor.202300154>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/lpor.202300154>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/lpor.202300154>.

# A 3D-RCAN for Image Restoration

A Manual for Using AI for Image Restoration

## A.1 Brief Introduction to the Model

Residual Channel Attention Networks ([RCAN](#) Zhang et al. [125]), are deep learning architectures initially designed for 2D super-resolution images, leveraging residual learning and channel attention mechanisms to prioritize high-frequency details. Their hierarchical “residual-in-residual” structure and adaptive channel-wise feature scaling enable superior recovery of fine textures in natural images. Researchers adapted RCAN for 3D fluorescence microscopy data to address challenges in volumetric imaging. This 3D-RCAN Chen et al. [21] implementation modifies the original architecture by replacing 2D [convolutions](#) with 3D operations, optimizing parameters (e.g., residual groups, channel dimensions) for GPU efficiency, and processing spatiotemporal (4D) data. This adaptation enables denoising and resolution enhancement of volumetric time-lapse microscopy, achieving prolonged live-cell imaging with minimal photo bleaching. By training on paired low/high-quality datasets (e.g., confocal  $\leftrightarrow$  STED or iSIM  $\leftrightarrow$  expansion microscopy), 3D-RCAN restores sub cellular structures, improves lateral and axial resolution, and outperforms existing networks (e.g., [CARE](#) Weigert et al. [117]) in fidelity metrics (e.g., SSIM Wang et al. [115] or peak signal to noise ratio). Its applications span fixed and live-cell studies, offering a versatile tool to overcome trade offs between spatiotemporal resolution and phototoxicity in fluorescence microscopy.

Several features of this neural network enable it to handle complex [image restoration](#) tasks effectively, as described in the paragraphs below. In later sections, we will employ this model as an example to show you how to undertake AI restoration in fluorescence microscopy.

### A.1.1 Residual Learning

3D-RCAN uses a residual learning approach, which means that the network learns the difference (or residual) between the input image and the desired output. Instead of learning the entire image transformation directly, the network focuses on predicting the fine details that need improvement. This makes it easier for the network to capture high-frequency details (like edges

and textures) while ignoring unnecessary background information. The network architecture is built using a residual-in-residual structure. This means that within each “residual block” (a unit of processing), there are even smaller “residual blocks.” This hierarchical design improves the network’s ability to learn intricate patterns in the data, allowing for better restoration of fine image details. The network essentially learns to focus on increasingly refined features at different levels of abstraction.

3D-RCAN incorporates skip connections in its architecture. Skip connections are shortcuts that allow the input image to bypass certain network layers and pass directly to later layers. There are long skip connections (which bypass multiple layers) and short skip connections (which skip only a few layers). These connections enable the network to focus on high-frequency information and avoid losing important details that may be “washed out” in deeper layers.

### A.1.2 Channel Attention Mechanism

One of the distinctive features of 3D-RCAN is its channel attention mechanism. In any image, different features (such as color or depth in the case of microscopy images) have different levels of influence in the image. The channel attention mechanism allows 3D-RCAN to focus on the most important channels by adaptively adjusting the weight given to each feature channel. This helps the network to enhance high-resolution details in relevant channels while minimizing the influence of less important features.

### A.1.3 3D Adaptation for Image Volumes

3D-RCAN is designed to handle not only 2D images but also 3D image volumes. In many microscopy techniques, images are captured as a series of 2D planes stacked together to form a 3D volume. RCAN adapts its structure to process these multi-dimensional data efficiently. By working in 3D, the network can improve both the lateral (XY) and axial (Z) resolution of the image volumes, which is especially important in applications like fluorescence microscopy.

### A.1.4 Channel Reduction for Efficient 3D Processing

By reducing channel dimensions before processing and restoring them afterward, the network minimizes redundant computations while retaining essential spatial and structural details. This mechanism enables deeper architectures without overwhelming GPU memory, particularly crucial for large 3D microscopy volumes. Coupled with channel attention, the compressed features are dynamically recalibrated to prioritize high-frequency details, ensuring robust restoration of fine structures like microtubules or mitochondrial membranes. The approach optimizes both training speed and inference performance, validated across synthetic and experimental datasets.

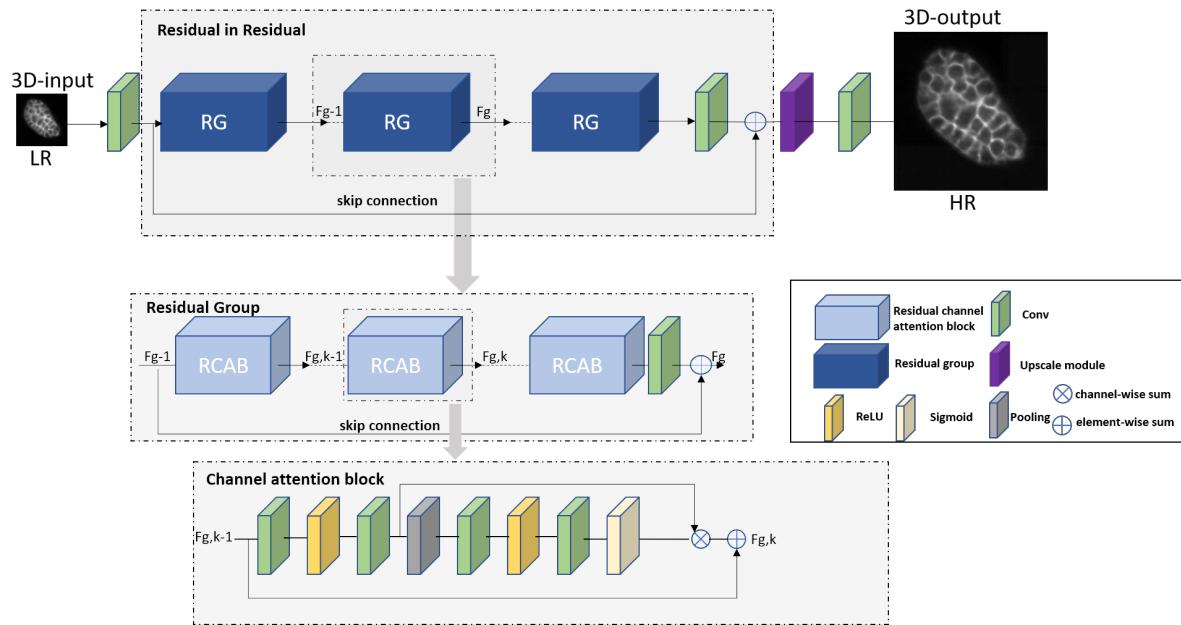


Figure A.1: **The architecture of 3D-RCAN.** This diagram illustrates the 3D Residual Channel Attention Network (RCAN) architecture for volumetric resolution enhancement. The low-resolution (LR) 3D input progresses through cascaded Residual Groups (RG), each containing Residual Channel Attention Blocks (RCAB) with channel attention to prioritize critical features. Skip connections preserve structural details across groups. The upscale module progressively enhances spatial resolution, culminating in the high-resolution (HR) 3D output.

## A.2 Overview of the Code and Data

### A.2.1 System Requirements

The basic system requirements are:

- Operating System: Windows 10, Linux, or Mac OS
- Python 3.7
- GPU: NVIDIA GPU with CUDA 10.0 and cuDNN 7.6.5

For more information about system requirements, please refer to the [3D-RCAN README](#).

### A.2.2 Installation of Dependencies

3D-RCAN itself does not require installation and installation of the required dependencies takes few seconds on a typical PC. The installation process involves three main steps:

1. **Check system requirements:** Ensure your system meets the basic requirements listed above, particularly Python 3.7 and CUDA compatibility.
2. **Create a new virtual environment:** Set up an isolated Python environment to avoid conflicts with other projects and ensure reproducibility.
3. **Install dependencies:** Download the repository and install required Python packages using the requirements.txt file provided in the [3D-RCAN repository](#).

### A.2.3 Quick Setup Steps

1. **Clone the repository:**

```
git clone https://github.com/AiviaCommunity/3D-RCAN.git
```

2. **Create and activate virtual environment:**

```
# Create virtual environment
python -m venv RCAN3D

# Activate (Windows)
.\RCAN3D\Scripts\activate

# Activate (macOS/Linux)
source RCAN3D/bin/activate
```

### 3. Install dependencies:

```
pip install -r requirements.txt
```

For comprehensive installation instructions, please refer to the *Dependencies Installation* section in the [3D-RCAN README](#).

#### A.2.4 Dataset

**Custom Data:** 3D-RCAN supports 3D data with strict low-quality/high-quality paired training (e.g., noisy/blurred inputs vs. high-SNR/clean GT) . Performance depends on domain alignment between training data (intensity distribution, resolution, noise characteristics) and target applications. The data should follow the format of 3D TIFF stacks (Z-Y-X order).

**Example datasets available:** You can leverage open-source datasets to obtain paired low-quality and high-quality data, such as the dataset from the 3D-RCAN paper. The 3D RCAN dataset is publicly available on [Zenodo](#). It includes paired low-quality and ground truth (GT) data for diverse applications: denoising multiple biological structures (actin, ER, Golgi, lysosomes, microtubules, and mitochondria), synthetic blurred phantom spheres with sharp GT, Confocal-to-STED microscopy modality transfer (microtubules, nuclear pores, DNA, and live-cell nuclei), and expansion microscopy (microtubules and mitochondria, pairing synthetic raw images with [deconvolved](#) iSIM images). Live-cell test data cover U2OS cells (mitochondria and lysosomes) and Jurkat T cells (EMTB-GFP). Raw data represent inputs (e.g., confocal images, noisy/blurred volumes), while GT corresponds to high-quality outputs (e.g., STED, deconvolved iSIM).

#### A.2.5 Training

Training a 3D-RCAN model requires a `config.json` file to specify parameters and data locations.

To initiate training, use the command:

```
python train.py -c config.json -o /path/to/training/output/dir
```

Training data can be specified in the `config.json` file by either:

1. Providing paths to folders containing raw and ground truth images (`training_data_dir`).
2. Listing specific pairs of raw and ground truth image files (`training_image_pairs`).

Numerous other options can be configured in the JSON file, including validation data, model architecture parameters (like `num_residual_blocks`, `num_residual_groups`), data augmentation settings, learning rate, loss functions, and metrics. The defaults for `num_residual_blocks` (3) and `num_residual_groups` (5) are set to balance performance and hardware constraints, aiming for optimal accuracy on standard GPUs (16-24GB VRAM) without causing memory overflow.

The expected runtime is approximately 5-10 minutes per epoch on a system similar to the tested environment (NVIDIA GeForce GTX 1080 Ti - 11GB) using the example `config.json`. Training progress and loss values can be monitored using TensorBoard.

For comprehensive details on configuring the `config.json` file, all available training options, and further instructions, please refer to the Training section in the [3D-RCAN README](#).

### A.2.6 Applying the Model

Trained 3D-RCAN models can be applied using the `apply.py` script. The script will automatically select the model with the lowest validation loss from the specified model directory.

There are two primary ways to apply a model:

1. To a single image

```
python apply.py -m /path/to/model_dir -i input_raw_image.tif -o output.tif
```

2. To a folder of images (batch mode)

```
python apply.py -m /path/to/model_dir  
-i /path/to/input_image_dir  
-o /path/to/output_image_dir
```

When processing a folder, ground truth images can also be specified using the `-g` argument for comparison.

Key command-line arguments include:

- `-m` or `--model_dir`: Path to the folder containing the trained model (required).
- `-i` or `--input`: Path to the input raw image or folder (required).
- `-o` or `--output`: Path for the output processed image or folder (required).
- `-g` or `--ground_truth`: Path to a reference ground truth image or folder (optional).
- `-b` or `--bpp`: Bit depth of the output image (e.g., 8, 16, 32).
- `-B` or `--block_shape`: Dimensions (Z,Y,X) for processing large images in blocks.
- `-O` or `--block_overlap_shape`: Overlap size (Z,Y,X) between blocks.
- `--normalize_output_range_between_zero_and_one`: Normalizes output intensity to the [0, 1] range, or to the full bit depth range (e.g., [0, 65535] for 16-bit) when combined with `-b`.

- **--rescale**: Performs affine rescaling to minimize MSE between restored and ground truth images, useful for comparisons similar to CARE methodology.
- **-f** or **--output\_tiff\_format**: Sets the output TIFF format (e.g., “imagej” or “ome”).

For a complete list of all options and detailed explanations, please refer to the Model Apply section in the [3D-RCAN README](#).

## A.3 Denoising Tutorial

### A.3.1 Data Preparation

**Input Format:** 3D TIFF stacks (Z-Y-X order)

**Data Pairs:** The [example image pairs](#) we employ are shown in Figure A.2 and are available in the denoising section of the [data for the 3D-RCAN paper](#).

- **Low-SNR Data:** Noisy images. Here we use low signal-to-noise ratio data obtained by imaging mitochondria labeled with mEmerald-Tomm20-C-10 in living U2OS cells using instant structured illumination microscopy under low illumination intensities.
- **High-SNR Ground Truth:** Corresponding clean images. Here we use high signal-to-noise ratio and high-resolution data obtained by imaging mitochondria labeled with mEmerald-Tomm20-C-10 in living U2OS cells using instant structured illumination microscopy under high illumination intensities.

**Directory Structure:** Organize the paired image data as follows.

```
dataset/
  train/
    raw/  # Training noisy data
    gt/   # Training clean data
  val/
    raw/  # Validation noisy data
    gt/   # Validation clean data
```

The validation data is not necessary for model training, but is an important part of assessing the model quality after training.

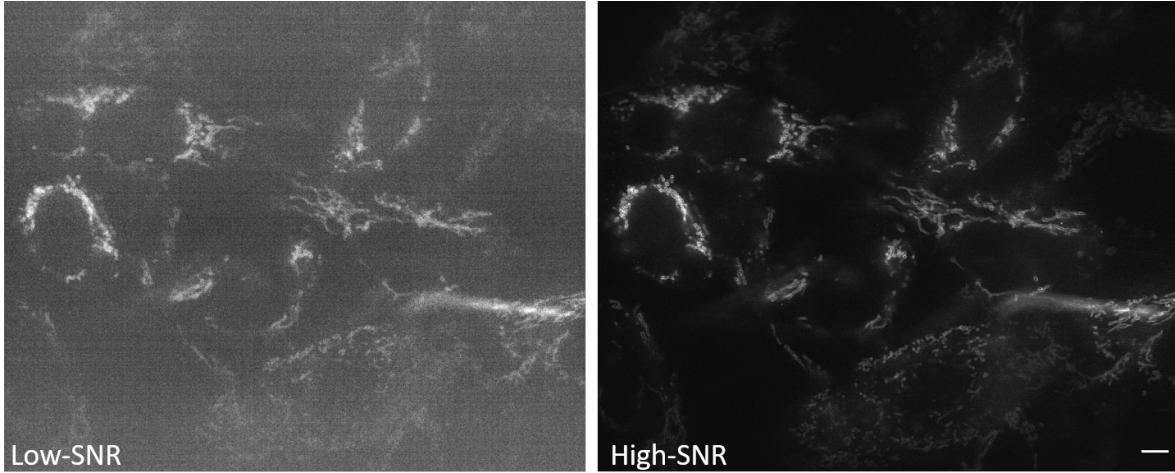


Figure A.2: **Dataset for denoising tasks.** (left) Low signal-to-noise ratio (SNR) image and (right) high-SNR image of mitochondria in U2OS cells. Images were acquired using instant structured illumination microscopy under low and high illumination intensities, respectively. Scale bar: 5 m.

### A.3.2 Training a Denoising Model

#### A.3.2.1 Configuration File

Configure the settings file (`config_denoise.json`) to define the data locations for the training/validation sets and the initial network hyperparameters. An example configuration file is shown below.

```
{
  "training_data_dir": {
    "raw": "E:\\Denoising\\Tomm20_Mitochondria\\Training\\Raw",
    "gt": "E:\\Denoising\\Tomm20_Mitochondria\\Training\\GT"
  },
  "validation_data_dir": {
    "raw": "E:\\Denoising\\Tomm20_Mitochondria\\Val\\raw",
    "gt": "E:\\Denoising\\Tomm20_Mitochondria\\Val\\gt"
  },
  "num_channels": 32,
  "num_residual_blocks": 3,
  "num_residual_groups": 5,
  "epochs": 100,
  "steps_per_epoch": 256,
```

```
    "initial_learning_rate": 1e-4  
}
```

### A.3.2.2 Training Command

Run the training using the following command, updating the path to where you have stored the example images as appropriate.

```
python train.py -c config_denoise.json  
-o "E:\\Denoising\\Tomm20_Mitochondria\\output"
```



Tip

If you are using a Mac or Linux system, replace the \ in the example paths with /.

### A.3.2.3 Training Outputs

The output directory will save the model parameters during the training process. For example, the file weights\_092\_0.06313289.hdf5 represents the model parameters saved at the 92nd training epoch, with a loss value of 0.06313289.

## A.3.3 Applying the Denoising Model

### A.3.3.1 Apply Command

To apply the model trained in the previous section, provide the model (-m), path for the input images (-i), and path for the output denoised images (-o).

```
python apply.py  
-m "E:\\Denoising\\Tomm20_Mitochondria\\output"  
-i "E:\\Denoising\\Tomm20_Mitochondria\\Test\\Raw"  
-o "E:\\Denoising\\Tomm20_Mitochondria\\Test\\Denoised"
```

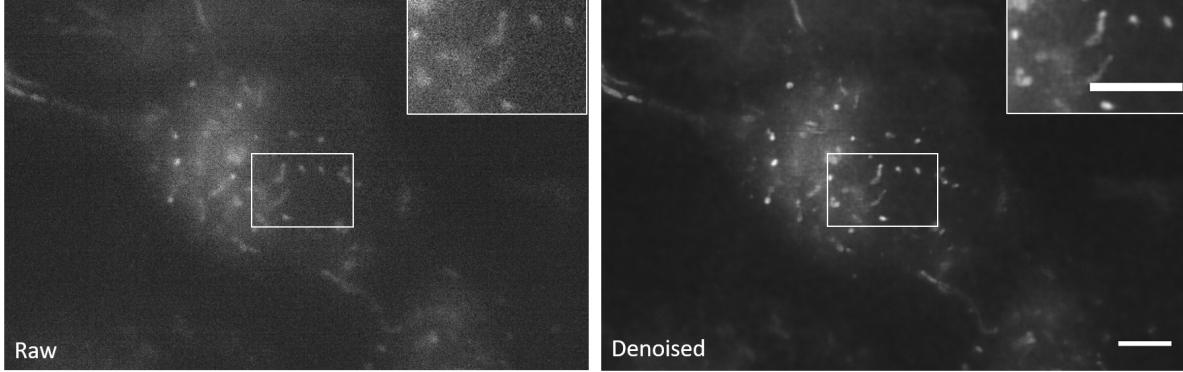


Figure A.3: **Denoising Performance.** The images show raw data input (left) and denoised image after using 3D-RCAN (right). Scale bars: 5 m for main image and 1 m for inset.

#### A.3.3.2 Output Analysis

Since the output TIFF file is a two-channel ImageJ Hyperstack containing raw and restored images, we can easily compare the denoising effects.

3D-RCAN reconstructs continuous organelle membranes (e.g., mitochondrial cristae) in noisy volumetric data by suppressing stochastic noise, achieving sub-diffraction structural integrity without amplifying high-frequency artifacts or distorting fine textures.

When evaluating the quality of denoised results generated by 3D-RCAN, PSNR (Peak Signal-to-Noise Ratio) and MSE (Mean Squared Error) serve as critical quantitative metrics. PSNR measures the logarithmic relationship between the maximum possible pixel intensity and the distortion introduced during restoration, with higher values indicating better alignment between the restored image and the ground truth (GT). Conversely, MSE calculates the average squared difference between corresponding pixels, where lower values reflect smaller pixel-level errors. For meaningful comparisons, the restored image should exhibit a higher PSNR and lower MSE relative to the raw input when both are evaluated against the GT. This directly quantifies the model's ability to enhance resolution while minimizing deviations from the true signal.

You can calculate PSNR and MSE with the formulas below, where  $m$ ,  $n$ , and  $p$  are the dimensions of the 3D image.  $I_{\text{raw}}(i, j, k)$  is the intensity value of the raw image at position  $(i, j, k)$  and  $I_{\text{GT}}(i, j, k)$  is the intensity value of the ground truth image at position  $(i, j, k)$ .

$$\text{MSE} = \frac{1}{mnp} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \sum_{k=0}^{p-1} [I_{\text{raw}}(i, j, k) - I_{\text{GT}}(i, j, k)]^2 \quad (\text{A.1})$$

$$\text{PSNR} = 10 \cdot \log_{10} \left( \frac{\text{MAX}_I^2}{\text{MSE}} \right) \quad (\text{A.2})$$

We recommend intensity normalization across raw, restored, and GT images (e.g., scaling to  $[0, 1]$ ) to avoid metric distortions. For heterogeneous samples, compute localized metrics within regions of interest (e.g., cell membranes vs. cytoplasm) to identify performance variations. Note that while high PSNR/MSE improvements generally correlate with perceptual quality, they should complement visual inspection and task-specific validations (e.g., downstream analysis accuracy). Always verify that the model’s output preserves biological relevance—metrics alone cannot capture context-specific artifacts or over-smoothing.

#### A.3.4 Further Denoising Guidance

To improve the performance of the 3D-RCAN model when unsatisfied with denoising results, we recommend starting by tuning hyperparameters in the `config.json` file. Start by adjusting the model architecture parameters to balance computational constraints and feature extraction capabilities. For instance, increasing `num_channels` (default:32) enhances the network’s capacity to learn complex features but requires more GPU memory. If limited by hardware, reducing it to 16-24 while moderately increasing `num_residual_groups` (default:5) or `num_residual_blocks` per group (default:3) could maintain depth without overwhelming resources. The `channel_reduction_ratio` (default:8) for attention mechanisms can be lowered to 4-6 to amplify the model’s sensitivity to channel-wise dependencies, though this should be validated using metrics to avoid overfitting.

Training dynamics can be optimized by revisiting the `initial_learning_rate` (default:1e-4). If training loss plateaus or fluctuates, gradually lowering it to 1e-5 stabilizes convergence, while aggressive learning rates (e.g., 5e-4) might accelerate early training but risk divergence. Switching the loss function from MAE (robust to noise) to MSE could prioritize pixel-level accuracy for high-frequency details, especially when ground truth data has sharp structures. Extending training epochs beyond the default 300—coupled with early stopping based on validation loss—helps the model capture subtle patterns in volumetric data.

Data-related parameters require careful calibration. Enabling `data_augmentation` (default:True) remains vital for small datasets to simulate variations in microscopy imaging, but it can be disabled for large, diverse datasets to speed up training. The `intensity_threshold` (default:0.25) and `area_ratio_threshold` (default:0.5) act as filters for low-quality patches. Raising the intensity threshold to 0.3-0.4 suppresses noisy backgrounds, while lowering the area ratio to 0.3-0.4 accommodates sparse biological structures like microtubules. For validation, always specify a dedicated `validation_data_dir` to monitor generalization and prevent overfitting.

Finally, hardware limitations can be mitigated by reducing `block_shape` dimensions during inference or simplifying the model architecture. Experiment iteratively: establish a baseline

with default settings, then progressively scale model complexity while tracking validation metrics like PSNR/SSIM. Consider hybrid strategies, such as combining MAE loss with SSIM regularization (a loss penalty using SSIM to preserve structural fidelity and prevent over-smoothing) or dynamic learning rate schedules, to adapt to specific data characteristics. Always validate changes against biologically relevant structures in your test set, as quantitative metrics alone may not reflect practical image quality for microscopy analysis.

## A.4 Deconvolution Tutorial

### A.4.1 Data Preparation

**Input Format:** 3D TIFF stacks (Z-Y-X order)

**Data Pairs:** The example image pairs we employ here are shown in Figure A.4 and are available in the ‘DL Decon’ section of [DeAbePlusData](#).

- **Input (Raw):** Low-quality, blurry 3D images with [PSF](#) such as single-view volumetric images acquired via diSPIM ( $0.8 \text{ NA} \times 0.8 \text{ NA}$  optics), with voxel size  $0.1625 \times 0.1625 \times 1.0 \mu\text{m}$
- **Ground Truth (GT):** High-quality, non-blurry 3D images such as dual-view joint deconvolution-reconstructed volumes (fusion of both diSPIM views), with isotropic resolution.

**Directory Structure:** Organize data as follows.

```
dataset/
  train/
    raw/  # Training raw data
    gt/   # Training label data
  val/
    raw/  # Validation raw data
    gt/   # Validation label data
```

The validation data is not necessary for model training, but is an important part of assessing the model quality after training.

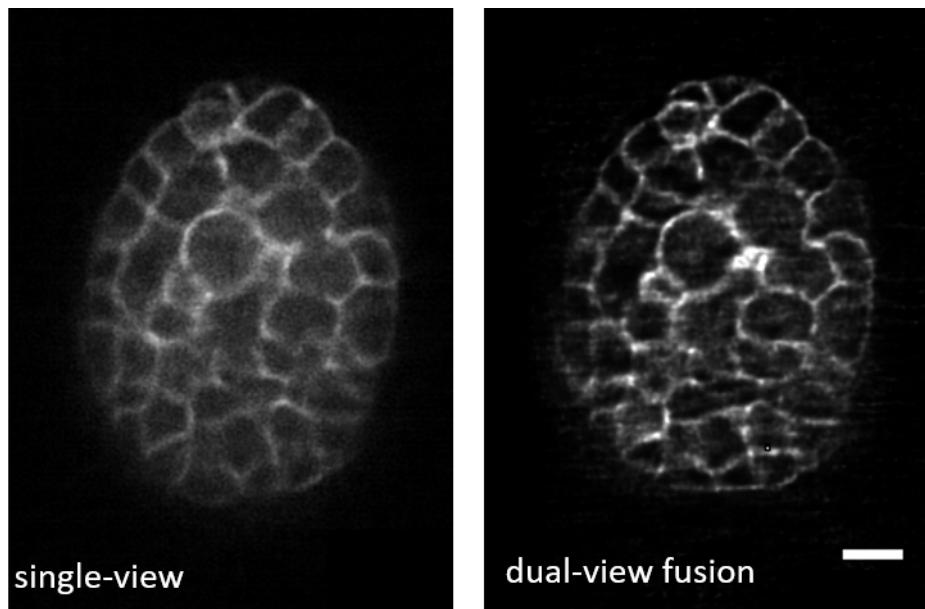


Figure A.4: **Dataset for Deconvolution.** The images show the single-view image (network input, left) and dual-view image (right) of *C. elegans* embryos. Scale bar: 5 m.

## A.4.2 Training a Deconvolution Model

### A.4.2.1 Configuration File

Configure the settings file (`config_decon.json`) to define the data locations for the training/validation sets and the initial network hyper-parameters.

```
{
  "training_data_dir": {
    "raw": "E:\\data_decon\\Decon_Training_Data\\Raw",
    "gt": "E:\\data_decon\\Decon_Training_Data\\gt"
  },
  "validation_data_dir": {
    "raw": "E:\\data_decon\\Decon_val_Data\\Raw",
    "gt": "E:\\data_decon\\Decon_val_Data\\gt"
  },
  "num_channels": 32,
  "num_residual_blocks": 3,
  "num_residual_groups": 5,
  "epochs": 200,
  "steps_per_epoch": 256,
```

```
    "initial_learning_rate": 1e-4  
}
```

#### A.4.2.2 Training Command

Run the training using the following command, updating the path to where you have stored the example images as appropriate.

```
python train.py -c config_decon.json -o "E:\\\\data_decon\\\\model_output"
```



Tip

If you are using a Mac or Linux system, replace the \ in the example paths with /.

#### A.4.2.3 Training Outputs:

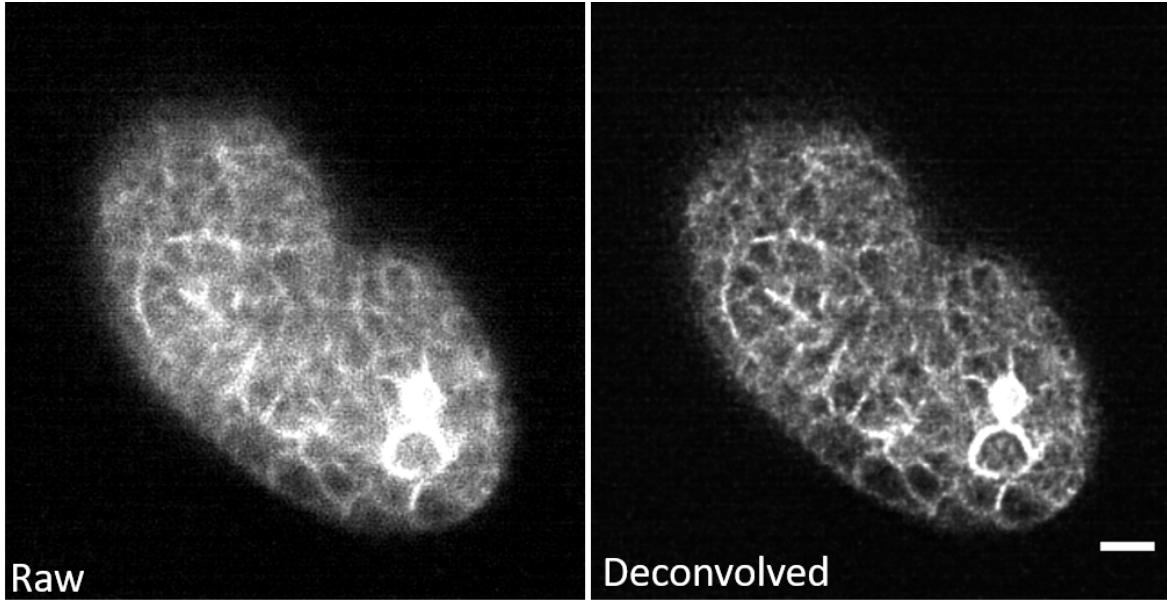
The output directory will save the model parameters during the training process. For example, the file weights\_092\_0.06313289.hdf5 represents the model parameters saved at the 92nd training epoch, with a loss value of 0.06313289.

### A.4.3 Applying the Deconvolution Model

#### A.4.3.1 Apply Command

To apply the model trained in the previous section, provide the model (-m), path for the input images (-i), and path for the output denoised images (-o).

```
python apply.py  
-m "E:\\\\data_decon\\\\model_output"  
-i "E:\\\\data_decon\\\\Decon_val_Data\\\\Raw"  
-o "E:\\\\data_decon\\\\Decon_val_Data\\\\deconvolved"
```



**Figure A.5: Network Prediction for Deconvolution.** The images show the raw data from *C. elegans* embryos (left) and 3D-RCAN prediction (right), demonstrating deconvolution. Scale bar: 5 m.

#### A.4.3.2 Output Analysis

Since the output TIFF file is a two-channel ImageJ Hyperstack containing raw and restored images, we can easily compare the deconvolved effects.

3D-RCAN corrects out-of-focus blur in GFP-labeled *C. elegans* embryos, resolving densely packed microtubule arrays within mitotic spindles and restoring cortical membrane contours obscured by optical scattering in thick embryonic samples, without introducing grid-like artifacts typical of Fourier-based deconvolution.

Refer to Section A.3.3.2 for baseline validation metrics. Compare with classical methods (e.g., Richardson Richardson [87] -Lucy Lucy [66] Deconvolution).

#### A.4.4 Further Deconvolution Guidance

Refer to Section A.3.4 for general guidance. For deconvolution, use a lightweight attention mechanism ( $\leq 3$  channel reduction steps) to avoid over-smoothing high-frequency structures.

## A.5 Aberration Correction Tutorial

### A.5.1 Data Preparation

**Input Format:** 3D TIFF stacks (Z-Y-X order)

**Data Pairs:** The sample image pairs employed here were *C. elegans* embryos expressing a GFP marker , as shown in Figure A.6. You can find these examples in the ‘DL DeAbe’ section of the [DeAbePlusData](#).

- **Aberrated Input Data:** Artificially degraded images, corrupted by synthetic aberrations modeled via [Zernike polynomials](#) (modes 1–7, defocus coefficient  $\leq 1.5$  rad, other modes  $\leq 0.5$  rad) and Poisson noise. Degradation is applied to shallow-layer ground truth using the physical forward model.
- **High-Quality Label Data:** Directly acquired shallow-layer images serve as ground truth, validated by their proximity to the objective lens (minimal inherent [aberrations](#)).

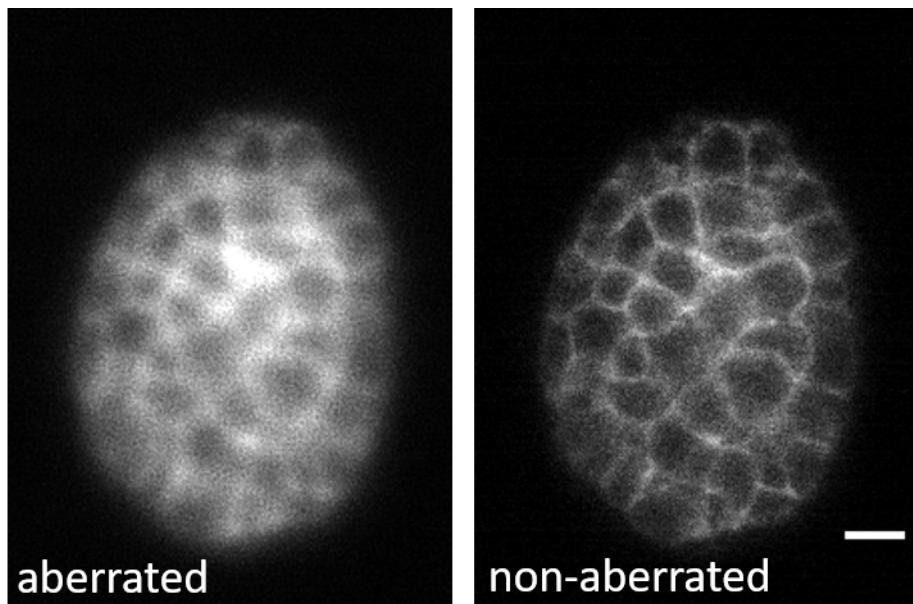


Figure A.6: **Dataset for Aberration Correction.** The images show the aberrated data (network input, left) and shallow-layer data (right) of *C. elegans* embryos expressing a GFP marker. Scale bar: 5 m.

**Directory Structure:** Organize data as follows.

```

dataset/
  train/
    raw/  # Training aberrated data
    gt/   # Training label data
  val/
    raw/  # Validation aberrated data
    gt/   # Validation label data

```

The validation data is not necessary for model training, but is an important part of assessing the model quality after training.

## A.5.2 Training an Aberration Correction Model

### A.5.2.1 Configuration File

Configure the settings file (`config_deabe.json`) to define the data locations for the training/validation sets and the initial network hyperparameters.

```

{
  "training_data_dir": {
    "raw": "E:\\data_aberration\\CropForTraining_ab\\Aberrated",
    "gt": "E:\\data_aberration\\CropForTraining_ab\\GT"
  },
  "validation_data_dir": {
    "raw": "E:\\data_aberration\\CropForTraining_ab\\Validation\\Aberrated",
    "gt": "E:\\data_aberration\\CropForTraining_ab\\Validation\\GT"
  },
  "num_channels": 32,
  "num_residual_blocks": 3,
  "num_residual_groups": 5,
  "epochs": 100,
  "steps_per_epoch": 100,
  "initial_learning_rate": 1e-4
}

```

### A.5.2.2 Training Command

Run the training using the following command, updating the path to where you have stored the example images as appropriate.

```
python train.py -c config_deabe.json  
-o "E:\\DeAbe\\Train\\output"
```



If you are using a Mac or Linux system, replace the \ in the example paths with /.

### A.5.2.3 Training Outputs:

The output directory will save the model parameters during the training process. For example, the file weights\_092\_0.06313289.hdf5 represents the model parameters saved at the 92nd training epoch, with a loss value of 0.06313289.

## A.5.3 Applying the Aberration Correction Model

### A.5.3.1 Apply Command

To apply the model trained in the previous section, provide the model (-m), path for the input images (-i), and path for the output denoised images (-o).

```
python apply.py  
-m "E:\\DeAbe\\Train\\output"  
-i "E:\\DeAbe\\Test\\Raw"  
-o "E:\\DeAbe\\Test\\Deaberrated"
```

### A.5.3.2 Output Analysis

Since the output TIFF file is a two-channel ImageJ Hyperstack containing raw and restored images, we can easily compare the de-aberration effects.

3D-RCAN effectively corrects system-induced optical aberrations in GFP-labeled *C. elegans* embryos. The de-aberrated image restores spatially ordered structures obscured by asymmetric blur in the raw data, while suppressing background granularity and enhancing signal-to-background contrast.

Refer to Section A.3.3.2 for baseline validation metrics. For aberration correction validation, analyze Fourier spectra to assess restored spatial symmetry and high-frequency energy, quantify wavefront errors via phase retrieval or bead-based PSF measurements, and evaluate intensity uniformity/edge sharpness across volumes.

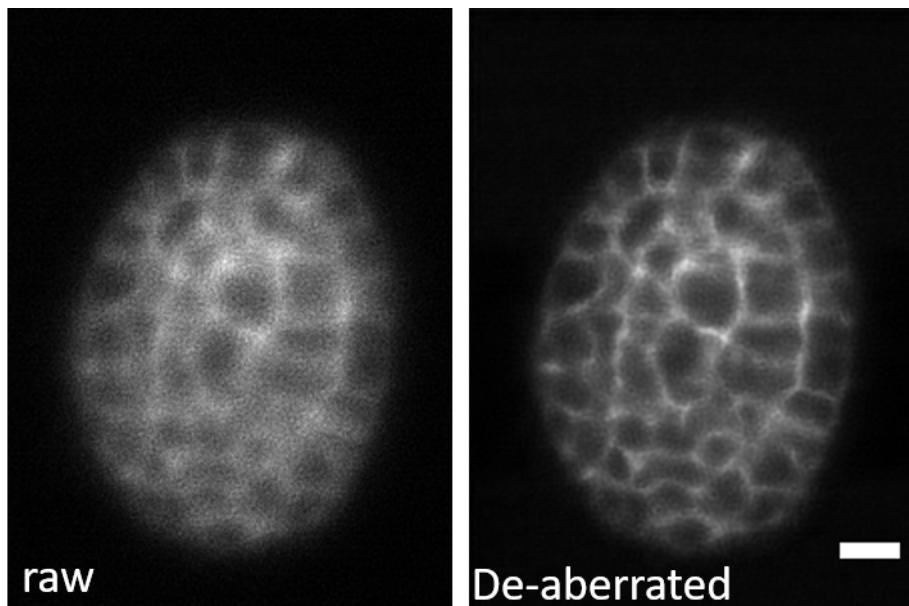


Figure A.7: **3D-RCAN Deaberration prediction.** The images show the raw data input to the network (left) and 3D-RCAN prediction (right), demonstrating aberration correction. Scale bar: 5 m.

#### A.5.4 Further Aberration Correction Guidance

Refer to Section A.3.4 for basic guidance. Prioritize RCAB layers with wider receptive fields (kernel size  $\leq 5^3$ ) to capture large-scale distortion patterns.

### A.6 Resolution Enhancement Tutorial

#### A.6.1 Data Preparation

**Input Format:** 3D TIFF stacks (Z-Y-X order)

**Data Pairs:** The [example image pairs](#) we employ are shown in Figure A.8. You can find these examples in the ‘confocal\_2\_STED’ section of the [data for the 3D-RCAN paper](#).

- **Raw data:** The “raw” data here refers to the original confocal microscopy images of the fixed mouse embryonic fibroblast (MEF) cells, specifically focusing on the microtubules. These images are acquired directly from the confocal microscope without any further processing or enhancement. They represent the initial, unaltered state of the microtubule structures as captured by the confocal microscope.

- **Ground Truth (GT) data:** The “ground truth” (GT) data consists of the high-resolution STED (Stimulated Emission Depletion) microscopy images of the same fixed MEF cells, again focusing on the microtubules.

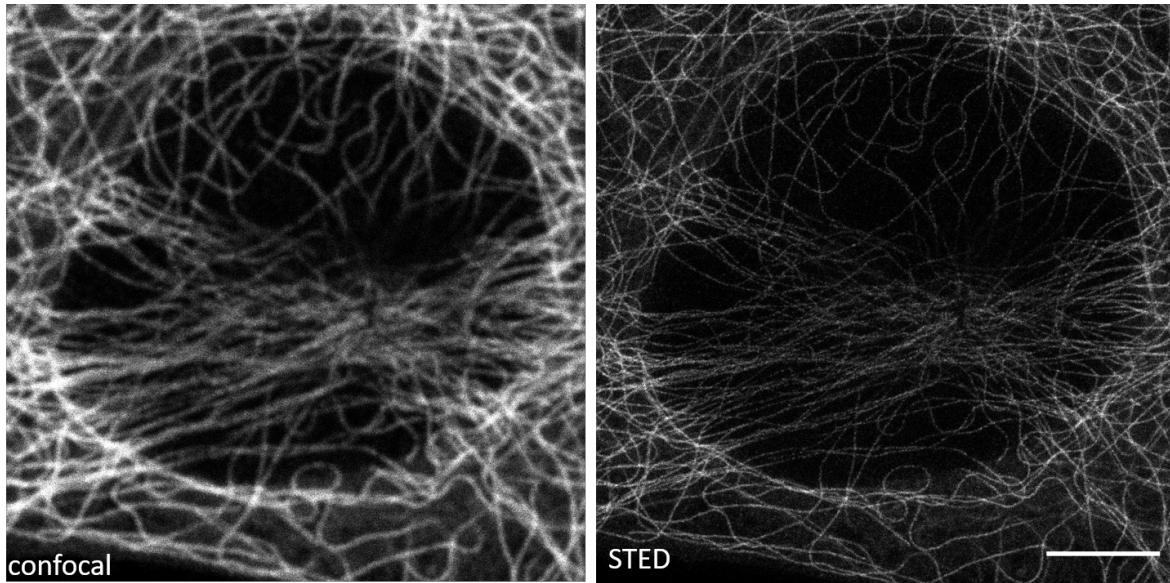


Figure A.8: **Dataset for resolution enhancement.** Image captured by confocal microscope (network input, left) and image captured by STED microscopy (ground truth, right). Both images show microtubules in mouse embryonic fibroblast cells. Scale bar: 5 m.

**Directory Structure:** Organize data as follows.

```
dataset/
  train/
    raw/  #confocal data
    gt/   #STED data
  val/
    raw/  # confocal data
    gt/   # STED data
```

The validation data is not necessary for model training, but is an important part of assessing the model quality after training.

## A.6.2 Training a Resolution Enhancement Model

### A.6.2.1 Configuration File

Configure the settings file (`config_sr.json`) to define the data locations for the training/validation sets and the initial network hyper-parameters.

```
{  
    "training_data_dir": {  
        "raw": "E:\\Confocal_2_STED\\Microtubule\\Training\\raw",  
        "gt": "E:\\Confocal_2_STED\\Microtubule\\Training\\gt1"  
    },  
  
    "num_channels": 32,  
    "num_residual_blocks": 3,  
    "num_residual_groups": 5,  
    "epochs": 100,  
    "steps_per_epoch": 256,  
    "initial_learning_rate": 1e-4  
}
```

### A.6.2.2 Training Command

Run the training using the following command, updating the path to where you have stored the example images as appropriate.

```
python train.py  
-c config_ex.json  
-o "E:\\Confocal_2_STED\\Microtubule\\Training\\output"
```



If you are using a Mac or Linux system, replace the \ in the example paths with /.

### A.6.3 Applying the Resolution Enhancement Model

#### A.6.3.1 Apply Command

To apply the model trained in the previous section, provide the model (-m), path for the input images (-i), and path for the output denoised images (-o).

```
python apply.py  
-m "E:\\Confocal_2_STED\\Microtubule\\Training\\output"  
-i "E:\\DATA_for_3dRCAN\\Confocal_2_STED\\Microtubule\\test\\"  
-o "E:\\Confocal_2_STED\\Microtubule\\test\\output\\"
```

#### A.6.3.2 Output Analysis

Since the output TIFF file is a two-channel ImageJ Hyperstack containing raw and restored images, we can easily investigate the extent of resolution enhancement.

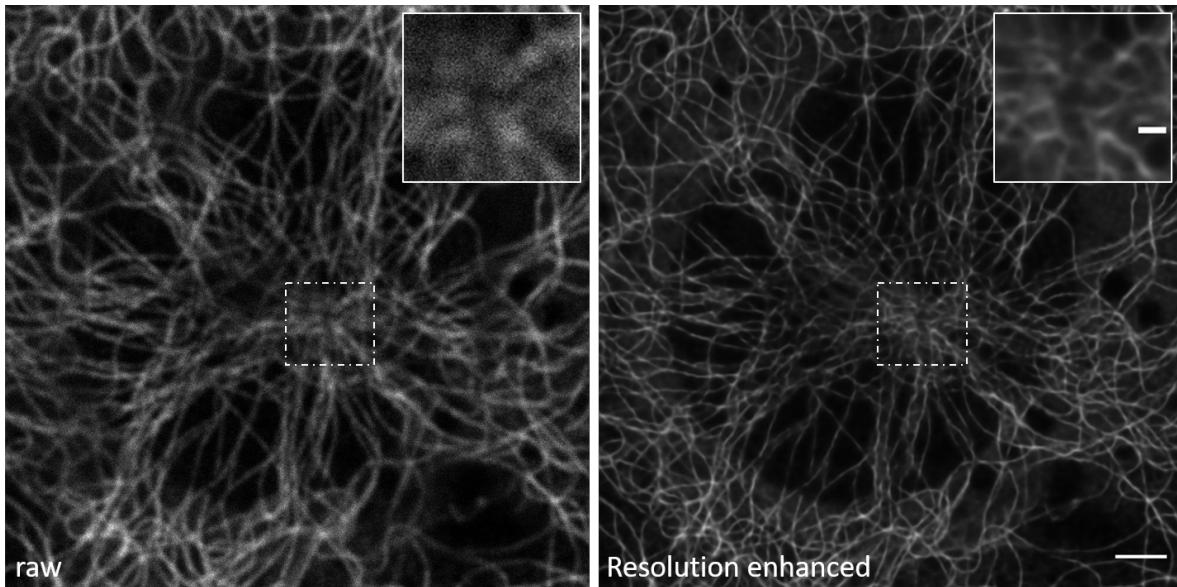


Figure A.9: **Network Prediction for Enhanced Resolution.** The images show the raw data of immunolabeled microtubules from mouse embryonic fibroblast cells (left) and 3D-RCAN predictions (right), demonstrating enhanced resolution. Scale bars: 5 μm for main figure and 1 μm for inset.

The model resolves densely packed microtubule intersections and eliminates axial blur in confocal-derived volumes, restoring filament topology to near-STED resolution.

Refer to Section [A.3.3.2](#) for baseline validation metrics. Check for artifacts in Z-stack transitions and consistency in organelle morphology.

#### **A.6.4 Further Resolution Enhancement Guidance**

Refer to Section [A.3.4](#) for basic guidance. Optimize residual groups (e.g., increase from 5 to 10) and channel dimensions ( $\leq 64$ ) to handle spatially variant upsampling.