AI IN SCIENCE TUTORIALS · WORKING WITH AI

SESSION 1

The Foundation

The Language of Seeing

Image analysis using Large Language Models — describing scientific images, running an analysis, and deciding whether to trust the result.

L. Fernando Santana UC Davis School of Medicine

A four-session series · L. Fernando Santana & Manuel F. Navedo

SOM Acceptable Use of Generative Artificial Intelligence in Academic Coursework Policy

THE SERIES

One series, four sessions, four guides

1

The Foundation

Santana

Describing images · running & trusting an analysis.

TODAY

2

Calcium Signals

Santana

Events in time — sparks, transients, kinetics.

3

Colocalization & Segmentation

Santana

Turning channels into objects and overlap.

4

Super-Resolution

Navedo

Localizations as point clouds · the capstone.

Today builds the shared foundation the other three sessions assume you already have.

LEARNING OBJECTIVES

By the end of this hour, you'll be able to…

01 **Describe** any scientific image in four dimensions — sample, imaging, structure, question.

02 **Judge** image quality — saturation, bit depth, sampling — before you measure anything.

03 **Drive** an analysis and read its overlay — without reading a line of code.

04 **Set** a threshold as a decision you own — and say how you chose it.

05 **Validate** a result against known ground truth before you trust it.

06 **Use discovery mode** to turn detected events into a real hypothesis, not just a new measurement.

THE THROUGH-LINE

Four principles that recur in every session

R

Reproducibility

Someone else can re-run it and get what you got.

A

Accountability

You own the result — the tool doesn't.

T

Transparency

The prompt and method are visible, not hidden.

V

Verification

You checked it against something you trust.

Today's anchor · Verification — you'll do it by hand before the hour is out, and every session returns to these four.

HOW EVERY SESSION WORKS

One arc, three moves — the spine of the series

01

Talk to the model

Describe, ask, explore — your first conversation, before any code. Execution or discovery.

→

02

Build the analysis

You and Claude develop the code that measures — or tests the hypothesis you just formed.

→

03

Validate it

Confirm the code is right — by its outputs and against known ground truth.

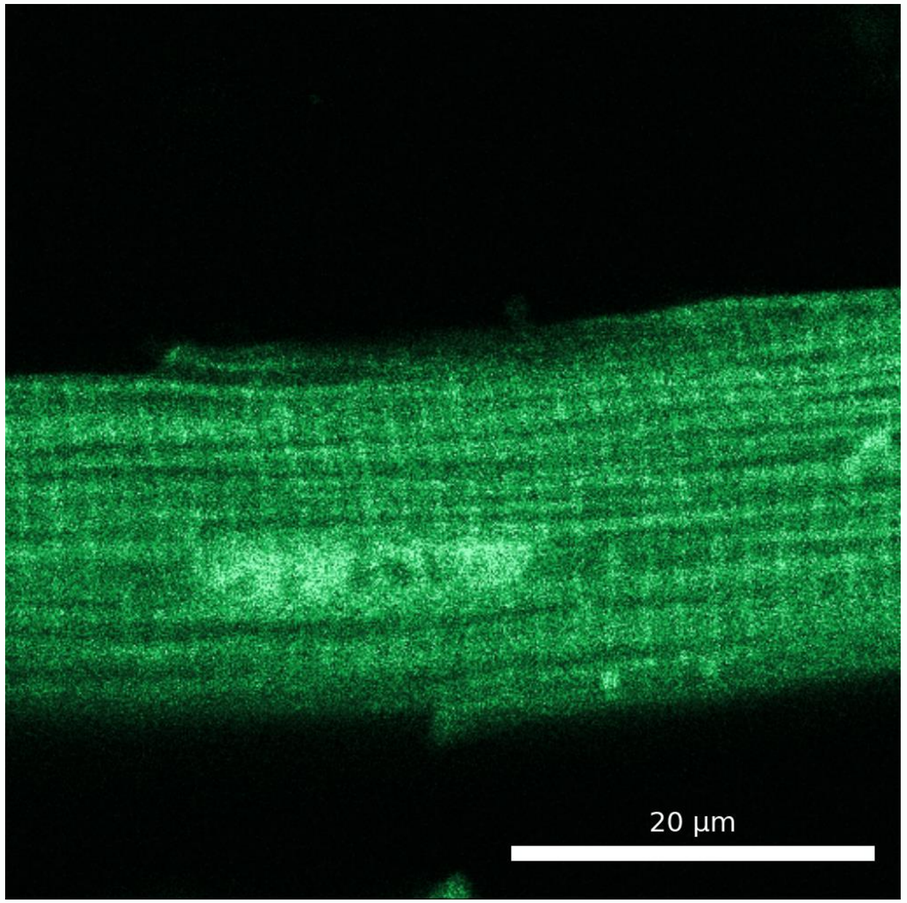
The modality changes each session. The arc doesn't.

01

The problem

Vague prompts don't fail loudly. They return confident, wrong answers — and look fine.

WATCH THIS The same image, two ways to ask



THE THIN PROMPT

"Analyze this image."

Claude *replying to "Analyze this image."*

This looks like a fluorescence microscopy image.
• An elongated, rod-shaped structure fills the frame — likely a labeled cell or fiber.
• Green signal — possibly GFP or Alexa 488.
• A regular banded pattern along the long axis; could be sarcomeres, or another periodic structure.
• A few darker rounded regions where the signal drops off.
• No obvious saturation.
I could measure length, width, intensity, band spacing, or the low-signal regions.

To give quantitative results I'd need more: what is the sample? What's the pixel size, or is there a scale bar? And what are you trying to measure?

It sounds competent — that's the trap. Under-briefed, the model hedges, guesses, and hands the work back to you.

WATCH THIS The rich prompt — same image, same model

THE RICH PROMPT

Adult mouse ventricular myocyte, cytosolic EGFP, 0.097 µm/px . A single nucleus shows as a bright spot; Z-lines as bright striations. Measure sarcomere spacing from the periodic bright Z-lines. Show the overlay first; expected spacing ~1.8–2.0 µm.

Claude · *rich prompt*

Pixel size 0.097 µm/px — read from metadata

Background 158.7 ± 9.6 → cutoff = mean + 3·SD = 187.4

Cell segmented — area ≈ 1,090 µm²

Sarcomere spacing 1.84 µm — within expected 1.8–2.0 µm ✓

THE KEY POINT

The gap between these two answers isn't what Claude can do.

It's what you gave it to work with.

The quality of the science you get out is proportional to the richness of the biological context you put in.

02

Describing what you see

Arc · step 1 of 3 — your first conversation. Four dimensions to cover, every time.

THE HABIT

Four dimensions of a scientific description

01

Sample

Organism, tissue, cell type, model, treatment.

02

Imaging

Modality, magnification, pixel size, channels, dyes.

03

Structure

Shapes, sizes, layout — expected and unexpected.

04

Question

The biological question, and what it's compared to.

Not a checklist to tick — a habit of thinking. Miss one and the analysis quietly drifts.

DIMENSION 01

Sample

Organism, tissue, cell type, disease model, treatment, timepoint — the biology Claude can't see, but you know.

THIN

"cardiomyocytes"

RICH

Adult mouse ventricular cardiomyocytes from a TAC heart-failure model , 16 weeks post-surgery.

DIMENSION 02

Imaging — five things to name

These five reappear in every session — learn them once, here.

01

Pixel = a measurement

Each pixel is a number — a light intensity, not just a color.

02

Bit depth

8-bit = 0–255; 16-bit = 0–65,535. Sets how fine the steps are.

03

Saturation

A pixel pinned at max is clipped — its true value is lost.

04

Spatial calibration

µm per pixel (nm in super-res). Turns pixels into real distance.

05

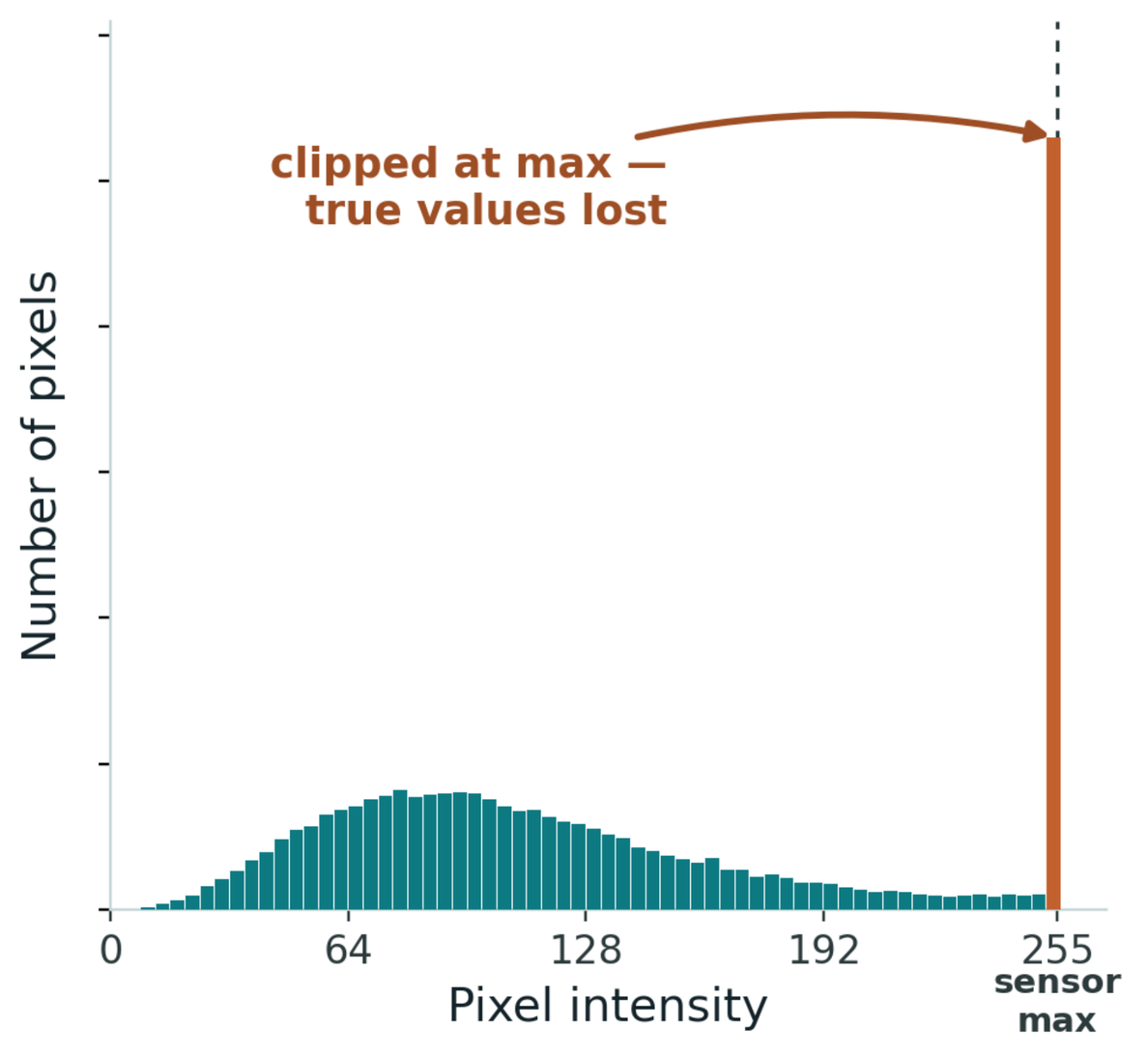
Channels

Which dye marks which structure — the heart of colocalization.

Four are quick. One bites. Saturation gets its own slide next — it's the one that silently destroys data.

DIMENSION 02 · THE ONE THAT BITES

Saturation — when the sensor stops counting



When a pixel hits the maximum, the sensor stopped counting. The real value might be 260 or 26,000 — you can't know. **No brightness, contrast, or analysis step recovers it afterward.**

THE HABIT

Check the histogram before you measure. A spike piled against the right edge means data is already lost — fix it at the microscope, not in analysis.

"Is anything in this image saturated?" is a fair first question to ask Claude.

DIMENSION 02 · ADEQUACY

Sampling — is your pixel small enough? (Nyquist)

Calibration tells you how big a pixel is. Nyquist tells you whether that size is small enough to capture the feature you care about.

Undersample → aliasing

Too few pixels per feature and fine or periodic detail is misread — you can even manufacture a periodicity that isn't there.

Oversample → waste

More light on the sample, faster bleaching, bigger files, slower runs — and no new information past the limit.

THE RULE OF THUMB

Put at least **2 pixels across the smallest detail** you want to resolve — about 2.3× in practice. Resolution is set by diffraction (≈ λ / 2·NA); your pixel should sit below half of it.

THIS IMAGE

0.097 µm/px, sarcomeres ~1.9 µm → **~19 pixels per band.** Well above Nyquist — which is why the Fourier spacing measurement is trustworthy.

HOMEWORK · BEFORE SESSION 2

Set the sampling

Work these through — the two examples, then one from your own work — and have Claude check each. You set the number; Claude verifies.

1 Spatial

Widefield, mCherry (λ ≈ 610 nm), NA 1.3. What's the finest pixel size worth using?

`d = λ / 2·NA, then pixel ≤ d / 2`

2 Temporal

A calcium transient peaks in ~40 ms. You want ≥ 8 points on the rise. What frame rate do you need?

`interval ≤ time ÷ points, then rate = 1 / interval`

Then, your own work: pick a structure or event you image, compute the pixel size or frame rate it needs — and have Claude check you, including whether you're oversampling.

DIMENSION 02 · THE FILE ITSELF

File structure — metadata & compression

A file is two things: the pixels, and a header recording what they mean — pixel size, bit depth, channels, objective, settings.

Metadata = your scale

Pixel size, channels and objective live in the file (.czi, .lif, .nd2, .oif). Have Claude read them — don't retype from memory. A PNG/JPEG export strips them; then you supply the pixel size.

Lossless is safe

PNG, or TIFF with LZW/ZIP, keeps every pixel value exact. Smaller files, no harm to measurement. Keep an uncompressed master.

Lossy corrupts data

JPEG discards information and alters pixel values to save space. Never quantify on a JPEG — the numbers are already changed.

The pattern: two silent corruptions bookend the work — saturation when you capture, lossy compression when you store. Both destroy the measurement before analysis begins.

DIMENSION 03

Structure

Shapes, sizes, spacing, organization — and what would be unexpected. Analogies help: *"like railroad ties," "like a cracked mosaic."*

THIN

"there are some bright spots"

RICH

Rod-shaped cells ~120 µm long with regular bright/dark banding at ~1.8 µm spacing — like evenly spaced rungs along each cell.

DIMENSION 04

Question

The specific biological question, why it matters, and what the answer is compared against — this steers every analytical choice.

THIN

"I want to count cells"

RICH

Does pressure overload change binucleation frequency? I'll compare TAC vs. sham , so I need a per-cell nucleus count , not a field total.

TRY IT NOW · 8 MIN

The Phone Call Test

Describe the image on screen as if you're on the phone with a computational biologist who has **never seen a heart cell** — and needs enough detail to design the analysis without seeing it.

1 Write · 5 min

Cover all four dimensions in plain prose. No prompt syntax yet.

2 Swap · 3 min

Trade with a neighbor. Could Claude act on theirs with no follow-up question?

WATCH FOR

Pixel size & sampling · file format · the biological question · artifacts · expected vs. unexpected.

TWO WAYS TO WORK

Execution & discovery — different prompts, different goals

Execution

you know what to measure

DESCRIBE — the rich four-dimension picture

SPECIFY — exact steps, output, "test on one first"

VERIFY — state how you'll check it, up front

Discovery

you want to know what's possible

SHOW — rich description, but don't direct the analysis

WONDER — "what could be measured that I haven't thought of?"

PROBE — chase what's interesting; ask for the biology

Both rest on the same rich description — which is why we taught it first. Today we drill execution.

03

Develop the analysis

Arc · step 2 of 3 — you and Claude build the code that measures, or tests your hypothesis.

TRY IT NOW · 10 MIN

Run your first execution prompt

PASTE THIS, THEN UPLOAD THE TIFF IMAGE + METADATA

```
Adult mouse ventricular myocyte, cytosolic EGFP (single channel).
Metadata is in cell_21_EGFP_metadata.txt (Olympus) — read the pixel
size from it; don't assume. Bright EGFP fills the cytoplasm; the
nucleus and the Z-lines appear bright.

1. Measure the background: from a cell-free region,
report mean and SD.

2. Segment the cell: outline where EGFP > mean + 3×SD.

3. Sarcomere spacing from the periodic bright Z-lines, in µm,
using the metadata's pixel size.

Show the overlay first, and report the background mean, SD, and cutoff
you used. Expected spacing ~1.8–2.0 µm — flag anything outside.
```

DO ONLY THIS

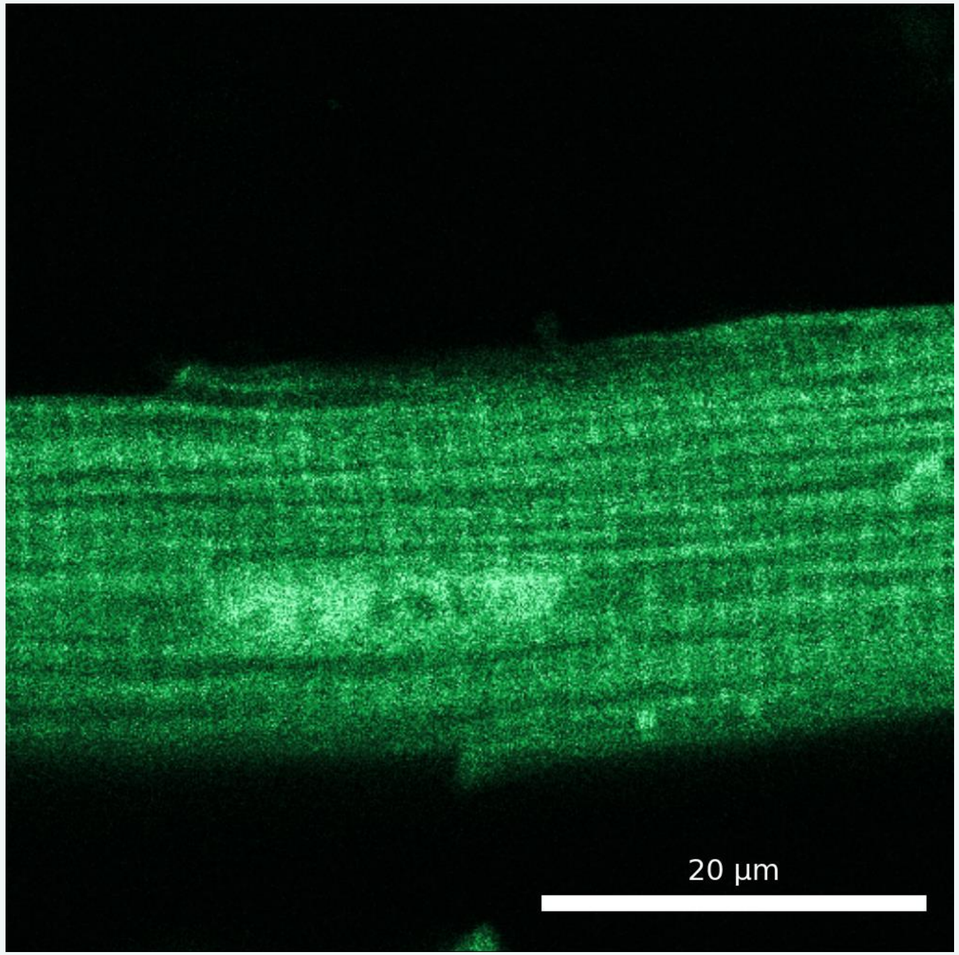
→ Run it.

→ Look at the overlay — nothing else yet.

→ Ask yourself one question:

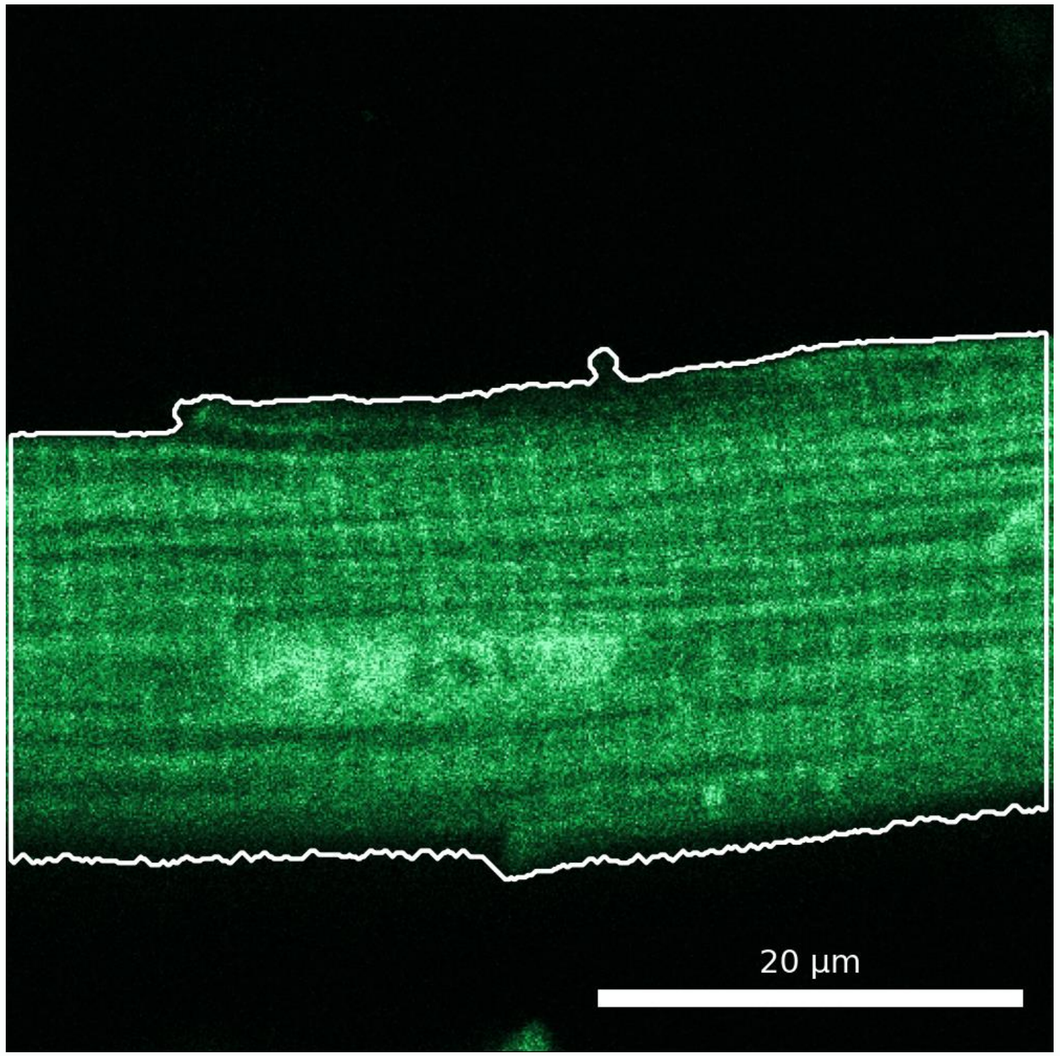
Did it outline the cells a biologist would outline?

Hold that thought — we'll turn it into a standard next.



THE RESULT YOU CAN READ

Reading the overlay



You don't have to read the code. You have to read the overlay — and that you can do.

If the overlay marks what a biologist would mark, the analysis is right — however the code is written.

If it doesn't, the analysis is wrong — however elegant the code looks.

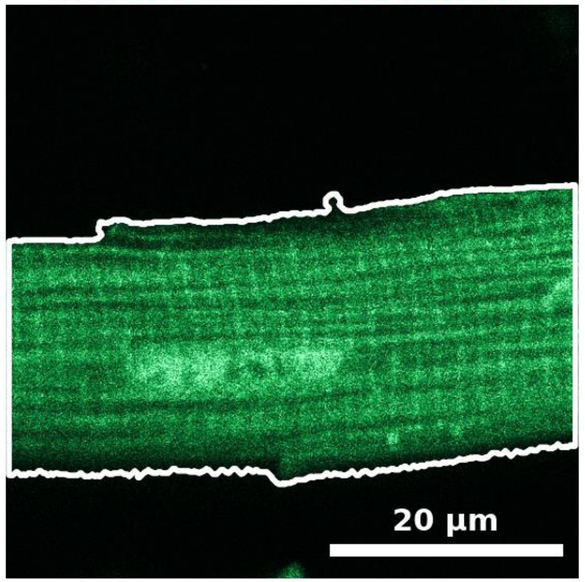
Code literacy is optional here. Biological judgment is not.

THE HIGHEST-LEVERAGE IDEA TODAY

Setting the threshold

A threshold is a decision you own, not a default you accept.

OUTPUT 1 · OVERLAY



OUTPUT 2 · REPORTED

Background 158.7 ± 9.6 → cutoff 187.4 (mean+3·SD)

Sarcomere spacing 1.84 µm — within 1.8–2.0 µm ✓

Turning intensities into objects means choosing a cut-off you own. The outline on the left is one cut's answer — move the cut, the outline moves.

```
SET IT FROM THE BACKGROUND, NOT BY EYE
cutoff = background mean + k × SD
choosing k — 3 is a common default — is the decision.
```

The same decision, three names ahead:

$\Delta F/F_0$ event cut · S2 **Segmentation cut · S3**

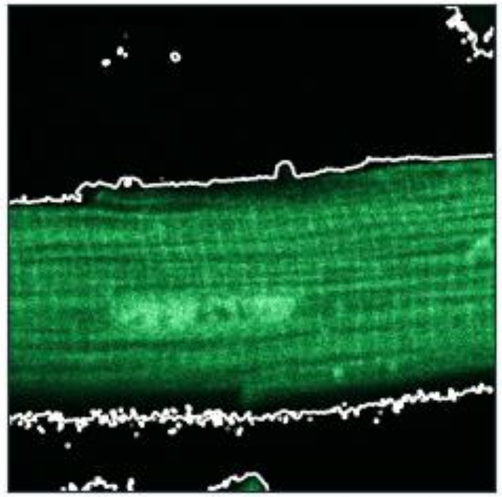
Localization cut · S4

EVERY CUT HAS A COST

The detection trade-off

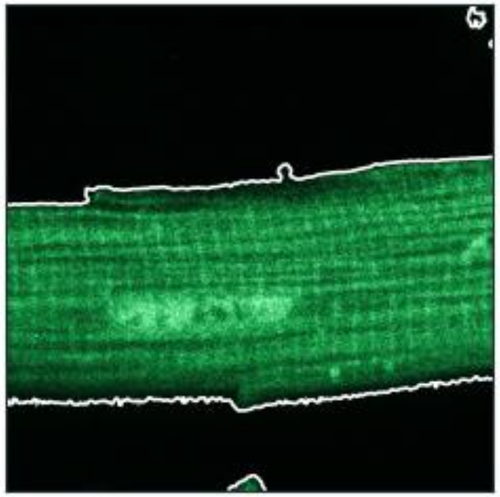
Watch the outline as the cut moves — you can't win both ways, so you choose where to live.

THRESHOLD TOO LOW



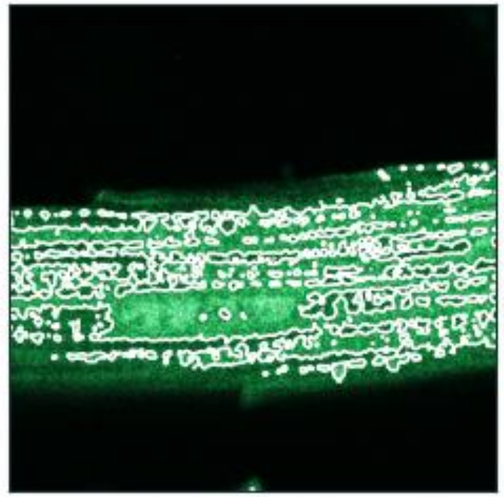
background detected as signal — false positives

JUST RIGHT



clean cell boundary — mean + 3·SD

THRESHOLD TOO HIGH



cell fragments, real signal lost — false negatives

Signal-to-noise sets the ceiling. The cleaner the separation between signal and background, the more room you have to choose well — low SNR, and every threshold hurts.

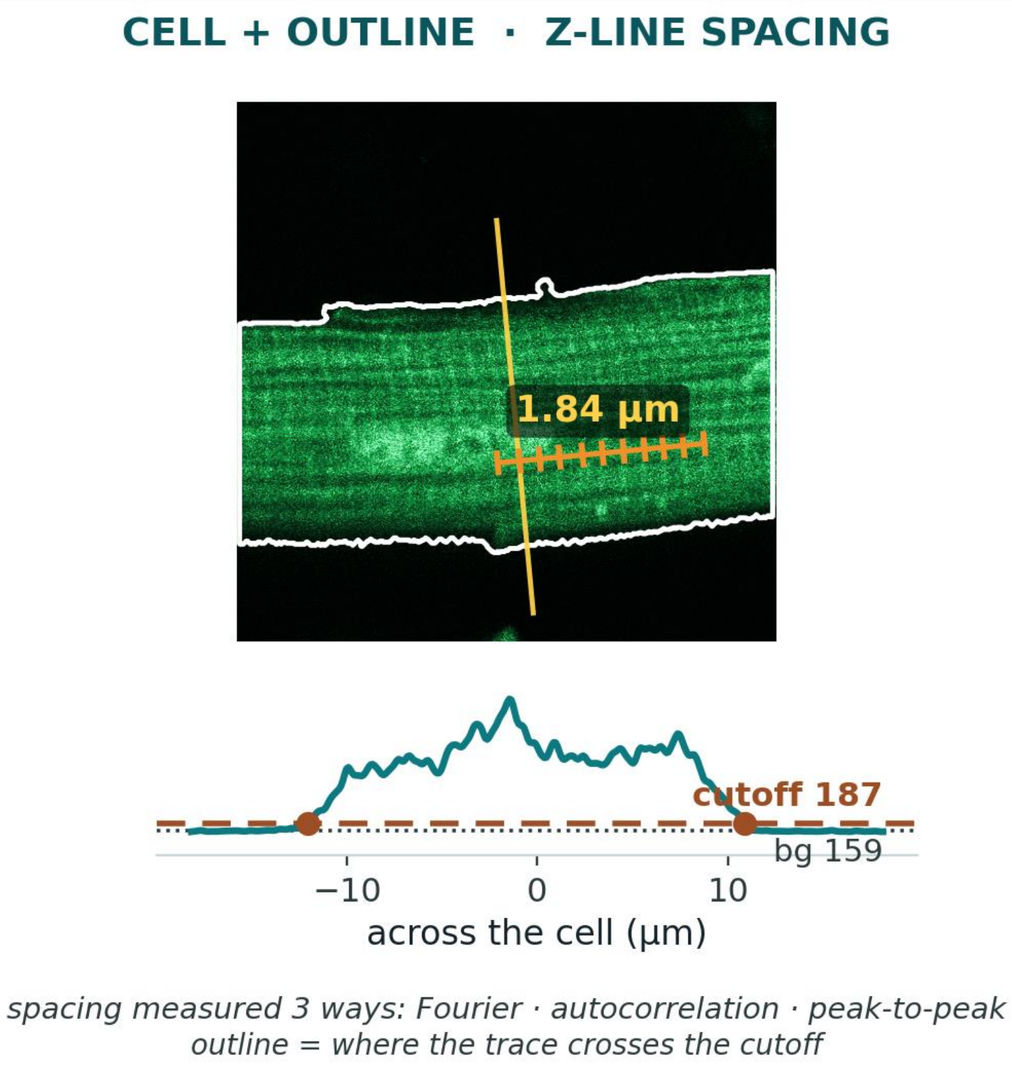
04

Validate before you trust

Arc · step 3 of 3 — confirm the code is right by its outputs, and against known ground truth.

BACK TO YOUR OVERLAY

So — why do you trust it?



Your overlay looked right. That's worth something. But **"looks right" is a feeling, not a measurement** — and confidently-wrong is the failure mode of this whole enterprise.

The honest answer: you don't trust it yet — not until you've checked it against something you already know.

VERIFY · MAKE IT JUSTIFY ITSELF

Ask it to check its own work

REVISED PROMPT

A follow-up to the analysis — turns “looks right” into “verified.”

“Before I trust the 1.84 µm, justify it and test it:”

1. **Explain each method**

   Fourier transform, autocorrelation, Z-line peak-to-peak — what each measures, and why it suits striated data.

2. **Show they agree**

   report the value each returns; if they differ, by how much, and which you’d trust.

3. **Validate against ground truth**

   generate a synthetic striated image with a known spacing, run this exact pipeline, and report the value it recovers.

4. **Name the failure modes**

   one condition, per method, under which it returns the wrong number on real data.

Transparency + verification + accountability, in one prompt.

You don't accept the number — you make the analysis defend it. This follow-up asks Claude to explain each method, cross-check them, test against a known answer, and admit where each breaks.

Transparency, verification, and accountability — the moment "looks right" becomes "verified."

VALIDATE · KNOW THE FAILURE MODES

What wrong looks like when it's confident

This tool fails *fluently* — the output reads like competence whether or not it's correct. Three shapes to recognize:

Fluent ≠ correct

Confident prose and clean-looking code are not evidence. Plausible is not verified.

Invented specifics

A method name, a citation, a parameter that sounds real and isn't. Check anything you can't place.

Random vs. systematic

3 wrong of 50, scattered → usually fine. 3 wrong the same way → systematic, and fixable with one instruction.

The one that bites is systematic error. Random misses hide inside biological variability; a systematic bias quietly shifts every number. The overlay and ground truth are how you tell them apart.

THE PROOF

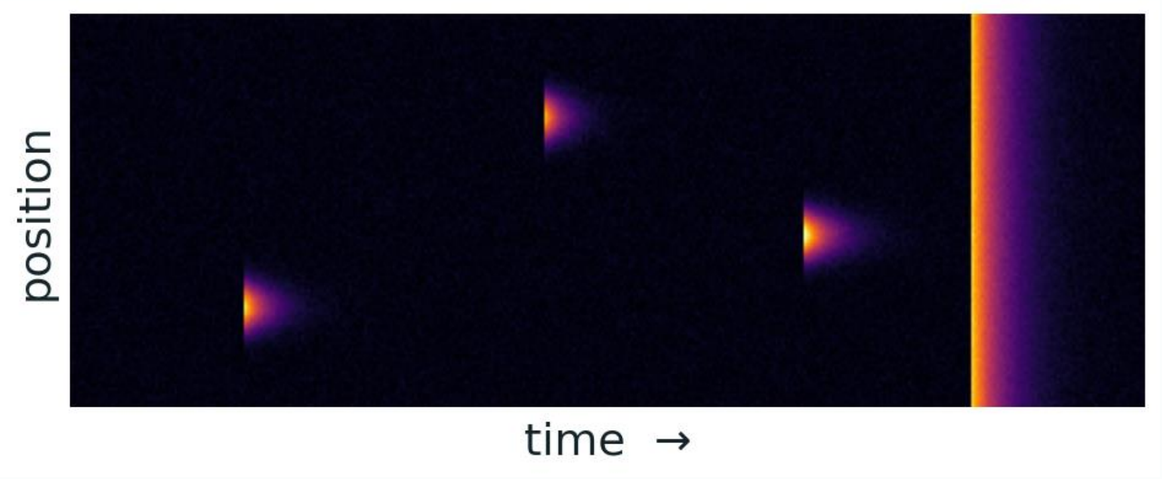
Known ground truth — test where you already know the answer

Build (or use) data with the answer planted in advance. Run the method on it. Count how much it recovers. **Every discrepancy is a diagnostic, not a mystery.**

1
Plant
Known features at known places

2
Run
The exact pipeline you'll use for real

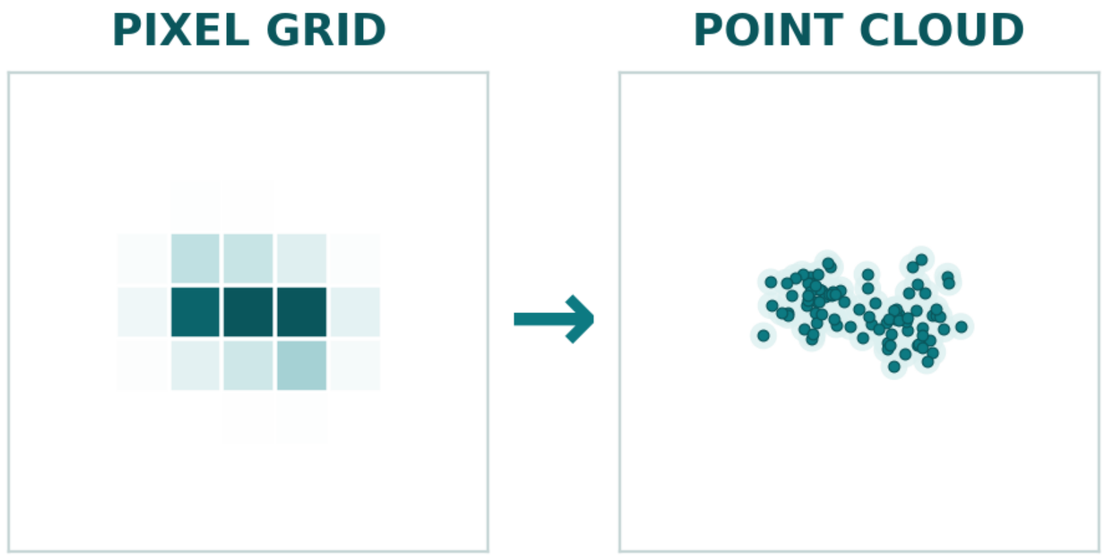
3
Compare
Recovered vs. truth — the gap is the lesson



You'll see this in its purest form in Session 2. The practice confocal linescan has exactly three calcium sparks and a whole-cell transient built in. Claude finds six, or eight? Now you know something concrete about your threshold — with zero biological ambiguity.

A SEAM WORTH NAMING

The pixel model — and the one place it breaks



Everything today is the same mental model: **an image is values on a grid.** Pixels, channels, thresholds, overlays — all of it lives there. Hold onto it; it carries you through Sessions 2 and 3.

Session 4 is where it stops being the right model. Super-resolution data isn't an image — it's a point cloud: a list of molecule positions, each with its own precision. Manuel hands you a new model when you need it.

Naming the seam now makes Session 4 a payoff, not a contradiction.

THE HANDOFF

What later sessions assume you walked in with

Walk into Session 2 with these six, and every session that follows can move fast — it references this, it doesn't re-teach it.

- ✓ Describe any image in four dimensions — sample, imaging, structure, question.
- ✓ Read pixel size, bit depth, saturation, channels — and judge sampling (Nyquist).
- ✓ Keep data quantitative — lossless formats, intact metadata, never JPEG.
- ✓ Treat a threshold as a decision you own, not a default to accept.
- ✓ Name the detection trade-off — and why SNR sets the ceiling.
- ✓ Validate against known ground truth before trusting a result.

THE OTHER MODE · SHOW → WONDER → PROBE

Discovery, live — ask what you haven't thought to measure

Claude · *discovery mode*

“What else could I measure — and how could I get a better image?”

MORE WAYS TO MEASURE

Periodicity, other ways
a 2D Fourier transform — the sarcomere spacing appears as a peak in frequency space (also autocorrelation, wavelets).

Nuclear shape
area and ellipticity (major:minor axis) of the dark nucleus; elongation tracks hypertrophy.

Sarcomere organization
spectral coherence (14× here) and orientation dispersion.

Texture entropy
disorder the mean intensity never captures.

WAYS TO IMPROVE THE IMAGE

Average more frames
line/frame averaging lifts SNR — sharpens every metric at left.

Use the headroom
128–1188 of 4095 — not saturated, so more gain or dwell is safe.

Widen the field
this cell runs off the frame — length and aspect need a lower-mag field.

Deconvolve
sharpen the Z-lines before measuring spacing.

Discovery isn’t only what to measure — it’s how to measure it better, and how to acquire it better. “I didn’t know I could ask that.”

You just executed — you knew what you wanted. Discovery flips it: set the scene, ask openly, and let Claude bring the analysis literature you don't have — both **what else to measure** and **how to image it better.**

1 · SHOW — Upload with the rich description — but don't say what to measure.

2 · WONDER — "What could I quantify that I haven't considered?"

3 · PROBE — Chase it: other ways to measure — and how to improve the image.

The central asymmetry: your biology, Claude's methods literature.

NEXT SESSION · 2

Calcium Signals

Confocal linescan calcium imaging — events in time.

BRING WITH YOU

The four-dimension description you wrote today — now add the temporal axis, and you're describing a calcium linescan.

Thank you — see you next session.