Seeing in the Dark: Intelligent Fourier Light Field Imaging for Bioluminescence Microscopy

Luis Felipe Morales Curiel and Michael Krieg*

Institut de Ciències Fotòniques, Castelldefels, The Barcelona Institute of Science and Technology, Barcelona, Spain

E-mail: michael.krieg@icfo.eu

Abstract

Bioluminescence microscopy offers a uniquely non-invasive window into cellular dynamics, yet its use has traditionally been limited by the intrinsically low brightness of luciferases. The poor photon budget forces long exposures, preventing faithful visualization of rapid physiological processes, especially in three dimensions. To overcome this barrier, we developed a Fourier light field microscope coupled with deep-learning–based reconstruction that achieves sub-second volumetric bioluminescence imaging with significantly improved spatial resolution. This approach eliminates the speed–resolution trade-off of conventional light field methods and bypasses the need for slow classical deconvolution. We demonstrate its power by performing real-time 3D calcium imaging in freely moving *Caenorhabditis elegans*, and by quantifying cell dynamics within stem-cell–derived spheroids using fluorescently labeled nuclei and calcium dynamics in muscles and neurons. Together, these results establish our framework as a practical tool for dynamic, volumetric studies of living systems.

Introduction

Seeing biological processes unfold deep inside living tissues, ideally label-free with no added light and minimal perturbation, is a longstanding goal in bioimaging. Autofluorescence of biological tissues and living cells has been used as a contrast agent,[1] but often requires ultraviolet excitation and prohibitively high illumination intensities due to the poor quantum yield of endogenous fluorophores.[2] Since the cloning of GFP, specific labels can be attached to individual biomolecules which afforded swift recording of biological dynamics in three dimensions in a confocal microscope. However, the need to introduce tags onto endogenous molecules or use fixed tissues, pose significant barriers to truly non-invasive imaging. An additional complication is the dependence on external illumination, which can generate spurious background signals from autofluorescence, introduce phototoxic effects,[3] and interfere with the intrinsic light sensitivity of biological specimens ranging from whole animals to subcellular organelles.

Bioluminescence offers an elegant solution through which proteins enzymatically produce light directly within the organism, without the need for an external excitation source.[4] However, the catalytic turnover of such enzymes, hence photon production, is slow such that bioluminescence imaging is traditionally constrained by the long exposure times required to collect enough photons with an acceptable signal-to-noise ratio.[5] Recently, we have shown that simple modification of the imaging protocols and technology and by combining the undersampled noisy images with content-aware restoration pipelines,[6] afforded a strong reduction exposure times down to 50 ms.[5] This balance of exposure and signal/noise ratio, accelerated by deep learning tech, was key to unlocking the second generation widefield bioluminescence microscopy. Intriguingly, the use of single-exposure light field imaging naturally complements bioluminescence by minimizing acquisition time and enabling 3D bioluminescence microscopy of subsecond calcium dynamics in freely moving animals.

Light field microscopy (LFM) improves this by capturing spatial and angular information using microlens arrays, enabling 3D reconstruction from a single exposure.[5,7,8] LFM is par-

ticularly well suited for photon-starved images, as a single exposure is able to record all the directional and spatial information of the imaged scene. This approach pushes the boundaries of what can be observed in living systems using excitation-free imaging modalities. Unfortunately, scanless 3-D imaging with LFM comes with a cost; lateral spatial resolution must be sacrificed to record angular information. Therefore, a trade-off between lateral and axial resolution is made, where the loss of lateral resolution is proportional to the number of discrete angular samples acquired.[9] This is because in standard LFM, spatial and angular information is directly mixed in the image plane, meaning that each microlens captures a different spatial region at multiple angles. This inherently reduces spatial resolution because pixels must be shared between these two types of information. A solution to this is Fourier light field microscopy (FLFM), in which the microlens array samples the light field in Fourier space,[10,11] meaning the angular spectrum (light ray direction) is now encoded within each microlens's aperture while sampling spatial details at higher frequencies. This affords diffraction-limited resolution in the x-y plane at an extended depth of field (DoF).

To visualize three-dimensional biological structures using LFM, raw image data must be computationally reconstructed into focused image planes. This involves disentangling spatial and angular light information captured by the sensor. A widely used iterative method to refine blurred images into sharper representations is the Richardson–Lucy deconvolution.[12–14] This approximation iteratively recovers a sharper estimate of the object by comparing the observed blurred image with the modeled or measured point spread function (PSF) of the imaging system. While this method can significantly enhance resolution, it is computationally intensive and often requires many iterations to converge to a high-fidelity reconstruction.

Artificial Intelligence (AI), particularly deep learning, has rapidly transformed the landscape of biological imaging.[15,16] By learning patterns from vast datasets, AI models can infer or reconstruct high-quality images from limited or noisy input,[6] offering a powerful alternative to traditional physics-based methods. In microscopy, AI has already enabled super-resolution reconstruction,[17] image denoising,[6] segmentation,[18] and rapid volumetric

imaging,[19] often surpassing the speed and accuracy of classical approaches.[5] These advances are especially impactful for techniques like LFM, where image formation is computationally demanding due to the complex mapping of angular and spatial information.[20] By leveraging AI, such systems can now decode and reconstruct 3D biological scenes in real time, even under ultra-low-light conditions, as is the case for bioluminescent signals.

Results and discussion

Design and performance of the Fourier light field microscope

We first established the parameters for the optical design, considering a target depth of field of ≈100um with a spatial resolution close to the diffraction limit and a field-of-view spanning 150 µm. We built the instrument around a 40×/1.25-NA silicon oil-immersion objective to maximize photon collection with a high numerical aperture avoiding spatial oversampling with an illumination path for fluorescence excitation and transmitted brightfield imaging. The optical parameters of the microscope were carefully selected to meet performance criteria as calculated by classical reconstruction using light-field deconvolution. We used a microlens array with a pitch of 1 mm and a focal length of 30.6 mm to target 2.1 μm lateral resolution (R_{xy}), 16.5 μm (R_z), 88 μm (DoF), and 147 μm (FOV). We positioned the MLA in the conjugated back focal plane, and projected its image onto a single photon resolving qCMOS camera. Additionally, an iris, positioned in the Fourier plane (conjugated to the back focal plane of the objective), acts as an angular filter that limits the range of light rays (angles) reaching the microlens array and sensor. Importantly, the iris radius in the Fourier plane corresponds to the angular acceptance defined by the system's NA (Fig. 1a).

To evaluate the performance of the FLF microscope we assessed the PSF by imaging d=200 nm fluorescent microspheres and measured the highest frequency distribution in the 2D power spectral density. In agreement with the design parameters and the target size, the largest recovered frequencies correspond to the smallest size in the image (Fig. 1b).

Importantly, each lenslet captures distinct information content, as shown by pairwise comparisons of structural similarity between the elemental images (Supplementary Fig. 1). We then tested the imaging procedure using gastrulating zebrafish embryos in short timelapse movies (Fig. 1d-f) but also in *C. elegans* (Fig. 1g) and spheroids generated from mouse embryonic stem cells (mESCs) (Fig. 1h). The latter two specimen were labeled in the nucleus with a fluorescent marker (GFP::lamin-1 (GFP::LMN-1) in the case of *C. elegans* and H2B-mTurquoise2 enhanced Nanolantern (H2B-TeNL) for mESC). We resolved individual nuclei and recorded frames every 3 min over a period of 27 min (Video S1). To restore the volumetric from the plenoptic information we resorted to conventional Richardson-Lucy deconvolution algorithms.[10] Although we were able to obtain high-quality 3D representation from plenoptic images (Fig. 1e), which facilitated single nuclei tracking in 3D over a volume of 120x120x100 µm (Fig. 1e), this classical deconvolution was computationally expensive and thus proved prohibitive for imaging longitudinal biological dynamics. Reconstructing the 92 perspectives of 3D data took 72 h to complete for 23 time points. In addition, artifacts in Richardson–Lucy light field deconvolution typically arise from noise amplification, inaccurate PSFs, or excessive iterations, leading to ringing, speckling, and instability in the reconstruction (Fig. 1g, h). To address this, we accelerated reconstruction with AI-based models.

Attention-based Deep Learning for Light Field Reconstruction

Supervised deep learning offers power and reliability but requires paired experimental and ground truth images,[21] obtained by simultaneous WF stack and plenoptic light field imaging. To enable multimodal imaging, we incorporate WF emission detection along the FLF path (Fig. 2a). Fluorescence excitation with WF detection provided paired ground truth data: a volumetric WF reconstruction and the corresponding plenoptic LF image of the same sample (Fig. 2a). This configuration enabled us to generate the paired WF and LF datasets required for supervised training.

Artificial neural networks have proven to be effective in accelerating volumetric reconstructions of LF microscopy, overcoming the computational bottlenecks of traditional deconvolution approaches.[5] However, standard convolutional neural networks (CNNs) process all spatial features uniformly due to their weight-sharing nature, which can lead to insufficient adaptation to local structures, potentially amplifying noise or blurring fine details in the reconstruction.[22] To address this, we developed and deployed a convolutional attention neural network optimized for Fourier light field (FLF) reconstruction.

Our model is built on a U-Net backbone and integrates two attention components: attention gates and convolutional attention blocks (Fig. 2b, ref[23]). Attention gates refine skip connections by selecting only the most relevant spatial information, while convolutional attention blocks, placed within the encoder, preserve high-frequency features typically lost during downsampling.[23] By integrating attention mechanisms, the network can dynamically weigh the importance of different spatial features, allowing it to focus on informative regions while suppressing irrelevant noise and artifacts.[24] This selective enhancement improves the fidelity of the reconstructed images by preserving fine structural details and reducing background interference.

Previously, we have shown that denoising input data before feeding them into the network improves reconstruction metrics.[5] Previous approaches required two separate networks: a content-aware restoration network[6] and a VCD network,[20] which increased the computational cost. In contrast, our end-to-end network is trained directly on noisy input projections and high-quality ground-truth volumes, enabling it to simultaneously learn image reconstruction and noise suppression. This behavior emerges from the structure of the data and the architecture design, rather than from explicit constraints or losses. As a result, the network performs reconstruction and noise attenuation simultaneously, without introducing additional complexity to the pipeline. Consequently, our model improves error and SSIM relative to models without denoising (Supplementary Fig. 2) and eliminates the need for preprocessing or auxiliary networks.

Furthermore, we used a custom loss function that combines root mean squared error (RMSE) with a Sobel-based edge loss to enhance spatial resolution and edge fidelity in the predicted output (Fig. 2c). While the MSE captures pixel-wise differences between the predicted and ground truth stacks, the Sobel component emphasizes discrepancies in edge information, helping the network better learn spatial structures. To further refine this, we applied the Sobel loss across all slices of the 3D stacks but introduced a Gaussian weighting scheme that prioritizes the central slices, the focal plane, where most of the relevant information is typically concentrated (Fig. 2c). This weighting ensures that errors in the focal region are penalized more heavily than those in peripheral slices, guiding the network to learn from the most informative regions of the data and resulting in reconstructions that match the ground-truth intensity distribution without introducing nonlinear distortions (Fig. 2d).

To train the network, we generated synthetic images by forward projection of high-resolution wide-field stacks using wave optics-based PSFs,[5,20] yielding $N = 2188$ pairs of images. While synthetic pre-training aided convergence, artifacts emerged from mismatches with real optics, leading to degraded performance (Fig. 2e). To enhance the network's relevance to real-world applications, we combined synthetic and experimental FLF data, in a three-stage strategy: pretraining with synthetic data, fine-tuning with experimental fluorescence images derived from *C. elegans* labeled with fluorescent protein in their muscles and neurons, and final training with bioluminescent data with the same expression pattern. We developed custom software to automatically extract 91 angular perspectives (200×200 pixels each) from raw FLF images (Fig. 2c). The pipeline supports varying microlens array configurations and aligns each input with its corresponding wide-field stack, acquired from the same sample using a motorized flip mirror system (Fig. 2a). The ground truth stacks were augmented via flipping, rotation, and z inversion to enhance robustness. Each WF volume spans 100 axial planes over a 100 µm depth and in total, we obtained $N = 2000$ paired datasets from our experiments. For each combination, we trained and evaluated

different models by measuring structural similarity (SSIM) and the RMSE (Fig. 2e, Table 1), and resolution by means of the power spectral density across reconstructions of unseen samples (Fig. 2h, Supplementary Fig. 4). In total, we compared 16 models (Table 1) and we named the best-performing model *LUCID* (*L*ightfield *U*nmixing and *C*omputational *I*mage *D*enoising). LUCID consistently suppressed noise, generalized across datasets (see validation section below), and preserved structural details (Fig. 2f, g Supplementary Fig. 6) with two orders of magnitude faster reconstruction times compared to classical Richardson Lucy deconvolution (Supplementary Fig. 5). Before applying the network to biological reconstructions, we analyzed how LUCID derives 3D volumes from raw FLF images using Grad-CAM to visualize model attention.[25–27]

Interpretability of the Fourier Light Field Reconstruction Network

Investigating Spatial Attention Patterns in the Reconstruction Network We examined attention maps in different stages of the U-Net: from the initial upsampling layers, through the encoder's second convolutional block, the central bottleneck, and the first decoder block. These maps reveal how the model progressively shifts from diffuse low-level attention in the early stages to more focused, structured regions that align with key features in the output volume (Fig. 3a).

The central "bottleneck" layer behaved differently: its attention maps were more abstract and less localized. This is expected as this layer compresses the entire image into a compact internal representation that captures overall structure and context, rather than fine detail. Importantly, the network design includes skip connections that pass spatial detail from earlier layers directly to later ones, allowing it to combine both global context and local image features.

These findings suggest that the network not only learns to reconstruct 3D images accurately but does so by focusing on biologically meaningful regions in a way that mirrors traditional image analysis - first parsing general features and then refining detail. Together,

this interpretability analysis fosters confidence and provides practical insight for evaluating emerging reconstruction methods.

Angular Attribution and View Selection Having shown that all views in principle carry distinct information content (Supplementary Fig.1), we aimed to understand now how different angular views contribute to volume reconstruction, we computed gradient-based attribution scores for each input channel (elemental image). Central views — near the middle of the FLF layout — consistently showed higher influence, while peripheral views contributed less, likely due to lower signal quality or redundancy (Fig. 3b).

To validate this, we performed *in silico* 'ablation experiments'. Removing the 5 most influential views led to a noticeable loss of fine detail, whereas removing 10 of the least important ones had minimal impact on reconstruction quality. This indicates that the network selectively leverages key perspectives for accurate 3D recovery and that attribution analysis can help optimize data acquisition, training pipelines and thus overall performance by identifying redundant views.

Validation

Cellular dynamics within mESC spheroids We first validated LUCID's performance by imaging spheroids derived from mouse embryonic stem cells with genetically encoded fluorescent nuclear markers. As before, all 92 lenslets in the microlens array capture unique information (Fig. 4a, b), and FLFM restoration using LUCID (Fig. 4c) was nearly real-time and took only 1 min for the whole video stack (t= 60 mins; 120 frames; 0.5 M pixels; Supplementary Video S), while classical deconvolution took 1.5 h to complete on the same workstation (Supplementary Fig. 5). In addition, the reconstruction has higher spatial resolution and structural similarity metrics (SSIM) computed against reference volumes compared to the deconvolved images (Fig. 4d).

We then performed 3D segmentation and calculated the proximity map, which indicates

the nearest cell neighbors (Fig. 4e) for every frame in the reconstructed image stack. We found that the average nearest-neighbor distance increases, concomitant with the growth of the cell cluster (Fig. 4f). This reveals how cells are packed in tissues and facilitates to study growth, (un)jamming transitions, cell sorting and migration, or tissue tension.

Next, we tracked each segmented nucleus (Fig. 4g) and calculated the mean squared displacement (Fig. 4f). Dimensional analysis revealed a scaling exponent larger than 1

$$\langle r^2(t) \rangle \sim t^\alpha \quad (\alpha > 1),$$

consistent with super-diffusive, directed motion rather than a random walk. This indicates that cells actively migrate apart, leading to reduced packing density and a transition of the spheroid from a more compact, solid-like state toward an expanding tissue during growth. Together, this shows that the single exposure 3D imaging reveal changes in tissue architecture in a nearly noninvasive fashion.

Imaging of nuclear dynamics in moving animals In conventional 3D time-lapse microscopy, capturing the full z-extent of a freely behaving animal requires multiple exposures that are fast enough to prevent motion artifacts and blurring. Because *C. elegans* have photoreceptors and avoid even low intensities of blue and orange light,[28,29] conventional fluorescence imaging causes frequent reversal, escape responses that obscure the natural behavioral repertoire.[30] Our bioluminescence light field approach addresses both challenges by enabling single-volume capture within a few hundred milliseconds, without external excitation light. To demonstrate this, we aim to resolve 3D tissue deformations in free-moving animals using the fluorescent marker strain that labels the nucleus(GFP::LMN-1) as fiducial markers. We used our initially trained LUCID model to restore short timelapse videos (Fig. 5a-c; Supplementary Video S) and segmented individual nuclei in these moving animals. We found that the network reconstruction using LUCID dramatically outperformed the RL deconvolution in terms of detected cells, their position (inset in Fig. 5d-e) and relia-

bility compared to ground truth, indicated by higher SSIM values (Fig. 5g). We then used short timelapse videos to segment and track individual nuclei in 3D of freely moving animals (Fig. 5h; Supplementary Video S). To obtain the local tissue deformation, we obtained the nearest-neighbor distance between selected nuclei (Fig. 5i), and calculated the change over time. This strain map indicates that cells are subject to both compression and extension, which brings nuclei closer or farther away from each other. Together, these results demonstrate that light field imaging combined with LUCID restoration enables quantitative strain mapping in freely moving animals, providing a powerful framework to link cellular-scale mechanics with tissue dynamics *in vivo*.

Excitation-free imaging of calcium dynamics Because bioluminescence microscopy relies on long exposure times, fast volumetric imaging of biological dynamics close to the diffraction limit has been out of reach so far. We recently showed that light-field microscopy marries well with photon-starved bioluminescent samples, due to their ability to capture volumetric information in one single exposure.[5] This approach reduces acquisition time by eliminating the need to scan through multiple planes, as required in conventional confocal or WF z-stacks, thereby preserving the frame rates necessary to capture rapid biological events such as calcium dynamics and stress responses. The trade-off, however, is reduced spatial resolution, which has so far limited our ability to record physiological signals from individual neurons.[5] Here, we extended this approach to reconstruct 3D image stacks in real time using bioluminescence Fourier light field microscopy.

To demonstrate the ability of our FLFM/LUCID combination to image photon-starved bioluminescent samples with short (≈1 s) exposure times, we first imaged partially constrained *C. elegans* expressing a calcium-sensitive NanoLantern in body wall muscles (Fig. 6a). As stated above, we trained the network directly with bioluminescence data. This adaptation aimed to enhance the network's ability to handle low-quality inputs while still producing high-quality reconstructions. Consequently, the 3D reconstruction of the biolu-

minescence scene was achieved seamlessly through this single-network pipeline, eliminating the need for a separate denoising network

As previously described, we observed that the bioluminescence intensity is correlated with the curvature during body bending,[5] indicative of higher calcium concentrations during muscle contractions (Fig. 6b-d).

To demonstrate that the FLFM is indeed capable of imaging at increased spatial resolution, we performed experiments using animals expressing a bioluminescent calcium indicator under the control of a panneuronal driver. We imaged the dorsorectal ganglion (Fig. 6e) of an immobilized animal and recorded spontaneous activation of a neuron that we identified as DVA,[31,32] and could clearly distinguish neighboring neurons that are 2 µm from one another (possibly DVC and DVB, Fig. 6f). After reconstruction, we found an isotropic resolution (Fig. 6g), which is higher than the ground truth presumably due to the suppression of noise and other artifacts (Fig. 6h).

Finally, we used our LUCID pipelines to reconstruct light-field images recorded from bioluminescent mESC derived spheroids. We compared the output between regular deconvolution and our network to the bioluminescent ground truth obtained from the regular z-stack by 3D wide-field imaging (Fig. 6 i-k). Again, we observed that LUCID performs better, as evaluated by the intensity distribution and the overall SSIM (Fig. 6l).

Together, these results establish bioluminescence FLFM, powered by our LUCID pipeline, as the first approach capable of capturing fast, volumetric bioluminescence signals at near-single-neuron resolution, overcoming the long-standing trade-off between sensitivity and spatiotemporal resolution in photon-starved samples.

Limitations

We noticed previously that the performance of the deep learning reconstructions is sensitive to input signal quality.[5] We therefore tested the ability of LUCID to reconstruct input images with varying SNR. We observed a clear performance drop below a specific, which

was quantified using SSIM against ground-truth data. For fluorescence imaging, the effective lower limit is SNR ≈ 2.2 (Supplementary Fig. 7), while for bioluminescence imaging it is approximately SNR ≈ 0.3 (Supplementary Fig. 8). Inputs below these thresholds result in poor reconstructions as the model cannot reliably distinguish signal from noise.

In our optical design, we prioritized the depth of field to better capture the volume of the sample. Increasing the number of microlenses along a given axis extends the depth of field by sampling more angular information, but this comes at the expense of lateral resolution because of reduced spatial sampling. In future designs, this trade-off between depth and resolution could be mitigated by reducing the number of lenslets, thereby improving spatial detail while still maintaining sufficient depth coverage. The consequence of this limitation can be seen in LMN-1::GFP-expressing *C. elegans*, where the nuclear envelope is poorly resolved because the intensity distribution is undersampled.

Our model also requires training on representative datasets to achieve optimal performance. Although transfer learning pipelines were implemented to ease this process, and we provide a ready-to-use Docker container, training remains a necessary step for domain adaptation to new biological samples or imaging conditions.

Conclusion

In this work, we developed and validated a Fourier light field microscope (FLFM) integrated with a deep learning reconstruction pipeline, LUCID, to overcome long-standing limitations in fast, volumetric imaging of photon-starved biological samples. Our optical design prioritized depth of field and field of view while maintaining near-diffraction-limited resolution, enabling single-exposure volumetric acquisition across diverse biological systems. Although classical deconvolution provided accurate reconstructions, it was prohibitively slow for longitudinal imaging. LUCID addressed this bottleneck by combining attention-based architectures, integrated denoising, and custom loss functions to achieve two orders of magnitude

faster reconstructions while preserving fine structural detail.

Interpretability analyses revealed that LUCID selectively leverages informative spatial features and angular perspectives, underscoring that the network learns biologically meaningful representations rather than hallucinating structure. Validation experiments across stem cell–derived spheroids and freely moving *C. elegans* demonstrated that LUCID-enabled FLFM provides isotropic 3D reconstructions at sufficient temporal resolution to quantify cell dynamics, tissue strain, and neuronal activity. Furthermore, we extended the framework to bioluminescence imaging, achieving near single-neuron resolution at short exposure times, thereby overcoming the traditional trade-off between sensitivity and spatiotemporal resolution in excitation-free microscopy.

Together, these results establish LUCID-enabled FLFM as a versatile and powerful platform for high-speed volumetric imaging under challenging low-SNR conditions. By bridging optical engineering with deep learning, our approach unlocks new opportunities to study fast biological dynamics, from tissue morphogenesis and mechanical strain to calcium signaling, in both fluorescent and bioluminescent systems. This work not only provides a technical foundation for future improvements in light field microscopy but also opens a path toward noninvasive, quantitative imaging of living systems with subcellular resolution.

Figures

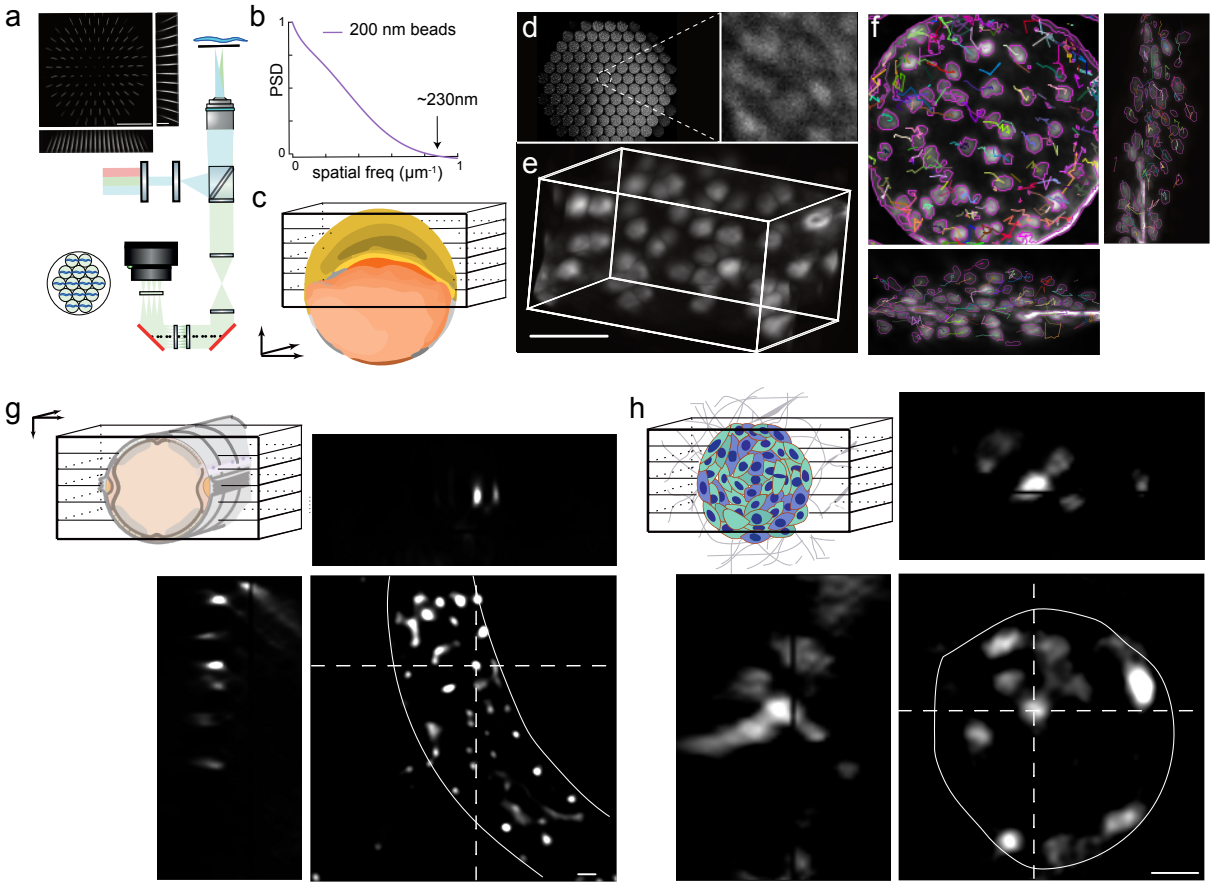
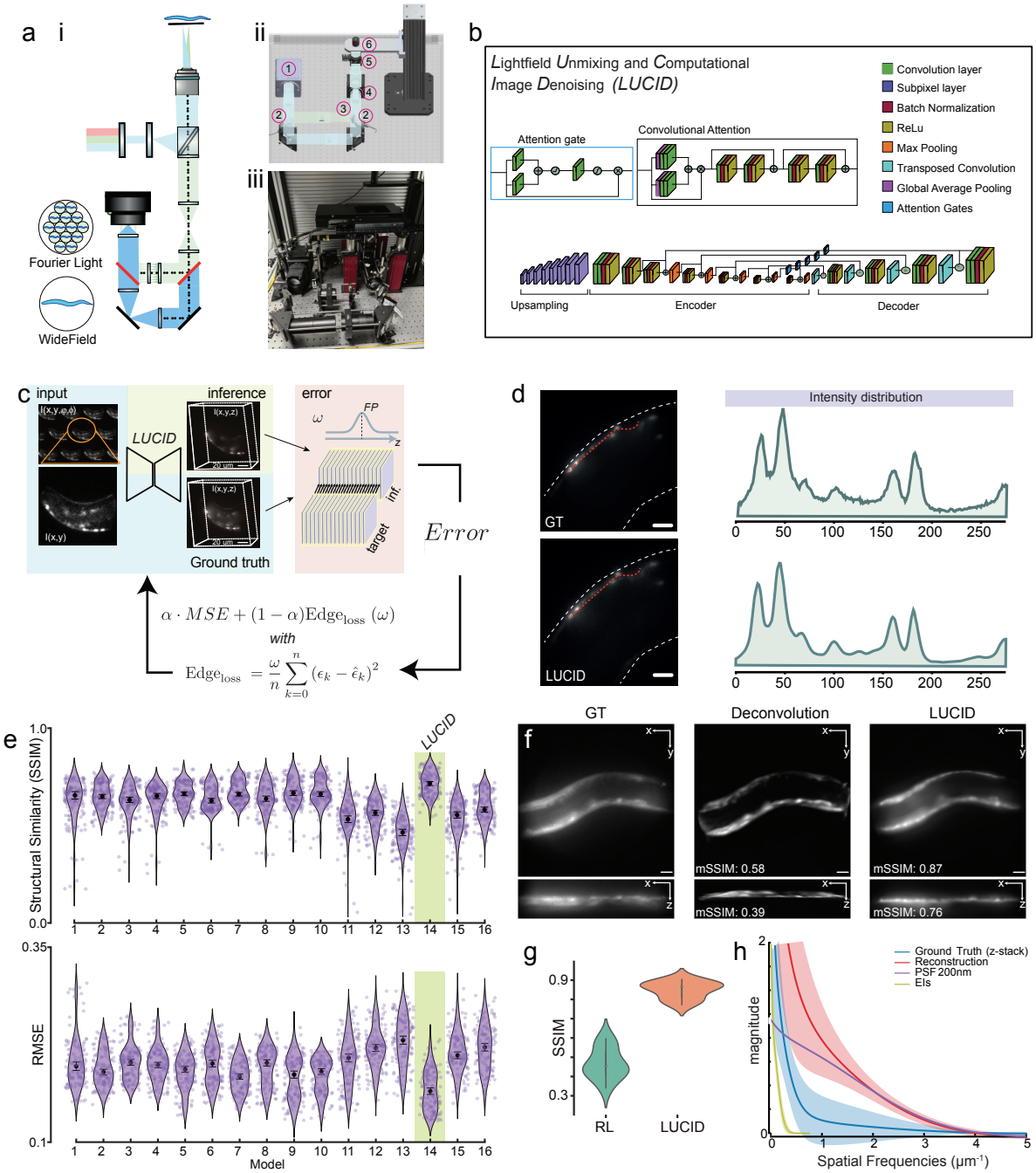


Figure 1: **Single-exposure Fourier light field microscopy. a.** Schematic of the Fourier light field microscope (FLFM). **b.** Spatial frequencies obtained from the images acquired with the FLFM, after Fourier transforming the point spread function (PSF). The upper inset shows the PSF of the system, experimentally measured with 200nm diameter fluorescent microspheres. **c.** Schematic of a zebrafish embryo at gastrulation stage. The square indicates the imaged region. **d.** Raw FLF image showing 92 perspectives of the zebrafish embryo through the microlens array, with the central elemental image magnified. Scale bar = 20μm. **e.** 3D reconstruction at time 0 from sequence shown in Supplementary Video using classical Richardson-Lucy deconvolution. Scale bar = 20μm. **f.** 3D segmentation of cells from a single plane of Supplementary Video , overlaid with their tracked trajectories across the X, Y, and Z projections. **g, h.** Classic Richardson-Lucy deconvolution of (g) *C. elegans* and (h) mouse stem cell colonies (5 days post-sampling) with fluorescently labeled nuclei. Images display the maximum intensity projection along the z-axis, with side views shown along the indicated dotted lines. Note the various artifacts in the reconstructions.

Figure 2 *(following page)*: **Deep learning pipelines for Multimodal Fourier light field microscopy. a.** i) Schematic of the dual optical-branch system for sequential acquisition of WF and FLF images. The setup uses motorized flipper mirrors to dynamically switch the optical path between two branches: one configured for FLF imaging and the other for conventional WF imaging. ii) 3D sketch of the setup. 1) Camera, 2) motorized flip-mirror, 3) MLA, 4) iris, 5) fluorescence path, 6) objective lense. iii) Actual photograph of the setup. **b.** Schematic of LUCID (Light-field unmixing and computational image denoising), an attention neural network for real-time volumetric bioluminescent imaging using an attention-based convolutional neural network (CNN) trained on synthetic and fluorescent data. **c.** Training pipeline for FLF models in 3D reconstruction. The process consists of five steps: (1) data acquisition, (2) extraction of EIs, (3) feedforward pass through LUCID, (4) comparison of predictions with the ground truth, and (5) gradient calculation to refine performance in the next cycle. A custom loss function combining Mean Squared Error (MSE) with a Sobel-based edge loss was used to improve spatial resolution and edge fidelity. The Sobel loss was applied across all slices of the 3D stacks with a Gaussian weighting scheme, ω, that emphasized the central, information-rich focal plane (FP, see Text and Methods for details). **d.** The comparison of intensity profiles between the GT and the neural network reconstruction along a line drawn on the edge of neurons in *C. elegans* (red dotted line), corresponds accurately to the presence or absence of neuronal signals. Input image SNR = 3.648. Reconstruction metrics mSSIM = 0.73, mNRMSE = 0.129. The pixels considered for the metrics are marked in the mask projection. Scale bar 50 μm. **e.** The training quality of the individual models was tested against unseen data by calculating the structural similarity (SSIM) and root mean squared error (RMSE) to the GT (N = 34 images/model, after augmentation N = 136 images/model). The best model, LUCID, with the lowest error and highest similarity was used to reconstruct experimental fluorescent images (p=0.0032 and p=0.001 compared to the 2nd (#) and 3rd ($*$) best models; Post Hoc Dunn Test with false-discovery rate p-adjustment). See Tab. 1 for details about models. **f.** Representative maximum intensity projection (upper panels) and side views (lower panels) of the ground truth,Richardson-Lucy (RL) deconvolution and AI reconstructions with LUCID of a worm expressing mTurquoise2 in the BWM. **g.** Reconstruction quality assessed through structural similarity comparing RL and LUCID performance. **h.** Fourier frequencies obtained after different steps of the image processing pipeline.


a i
Fourier Light
WideField
ii
iii
b
Lightfield Unmixing and Computational Image Denoising (LUCID)
Convolution layer
Subpixel layer
Batch Normalization
ReLu
Max Pooling
Transposed Convolution
Global Average Pooling
Attention Gates
Attention gate
Convolutional Attention
Upsampling
Encoder
Decoder
c
input
inference
error
LUCID
Ground truth
target
inf.
Error
$\alpha \cdot MSE + (1-\alpha)\mathrm{Edge}_{\mathrm{loss}}(\omega)$
with
$\mathrm{Edge}_{\mathrm{loss}} = \frac{\omega}{n}\sum_{k=0}^{n}(\epsilon_k - \hat{\epsilon}_k)^2$
d
GT
LUCID
Intensity distribution
e
Structural Similarity (SSIM)
RMSE
Model
LUCID
f
GT
Deconvolution
LUCID
mSSIM: 0.58
mSSIM: 0.87
mSSIM: 0.39
mSSIM: 0.76
g
SSIM
RL
LUCID
h
Ground Truth (z-stack)
Reconstruction
PSF 200nm
EIs
magnitude
Spatial Frequencies (μm⁻¹)

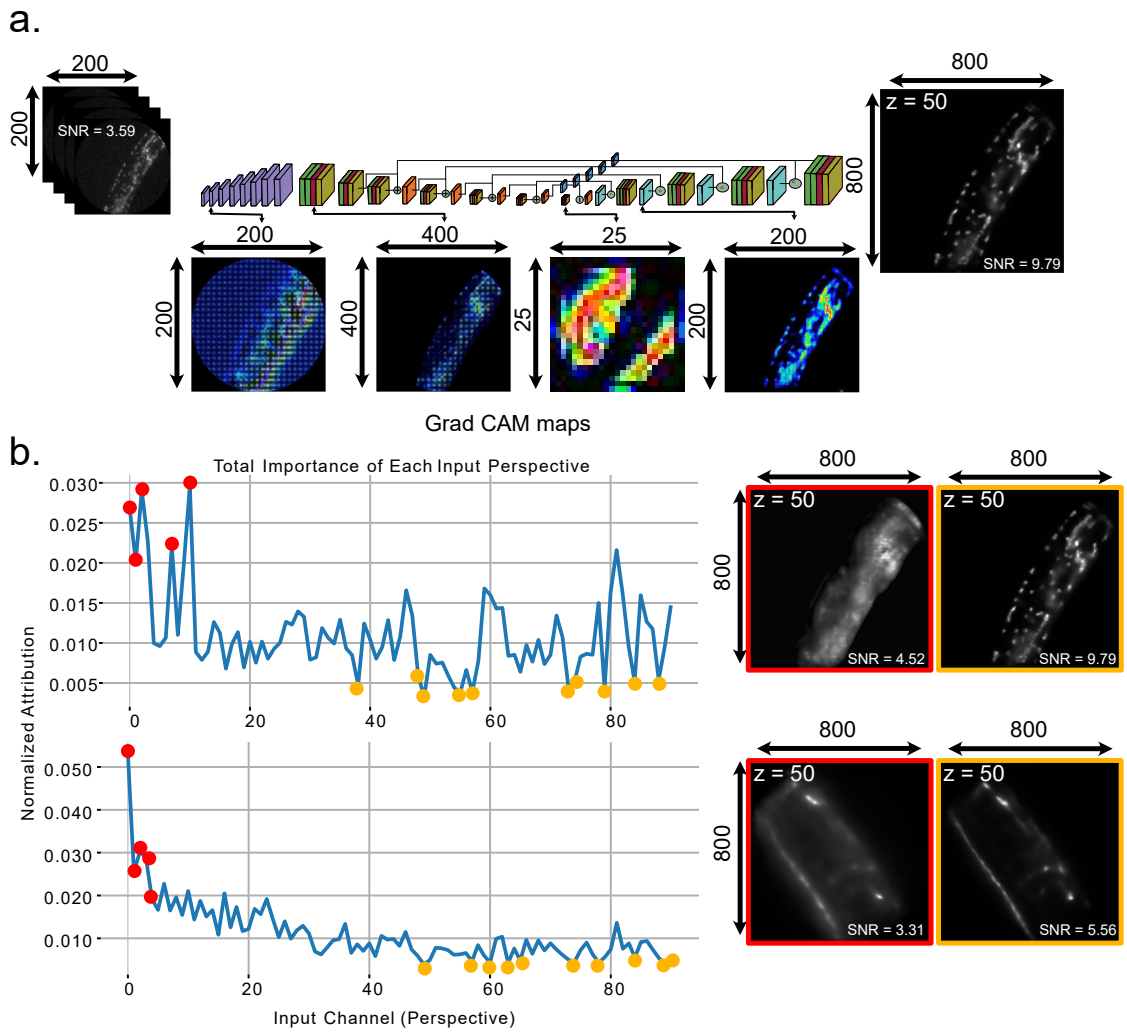


Figure 3: **Interpretability analysis of Fourier light field reconstruction network. a.** Grad-CAM spatial attention maps extracted at various stages of the U-Net, including input, bottleneck, and decoder layers, reveal how the network progressively learns to attend to salient spatial regions in the input. Heatmaps correspond to the reconstruction of the output slice at $z = 50$, the most information-rich plane. While input and decoder stage maps clearly highlight object structure, the bottleneck layer exhibits diffuse and spatially abstracted activations due to latent compression and feature generalization. Images depicted correspond to the *C. elegans* GFP::*lmn-1* strain. **b.** Channel-wise attribution analysis showing the relative contribution of each input perspective (elemental image) to the reconstruction, based on gradient magnitude. Two different networks trained on distinct datasets: fluorescence nuclear data (upper panels) and bioluminescence *rab-3* neuron data (lower panels). Each with unique attribution profiles. Red and yellow markers indicate the five most and ten least influential views, respectively. Reconstructions with red and yellow frames correspond to versions where these important/unimportant views were ablated. Removal of the most influential perspectives leads to severe loss of high-frequency content, while removal of the least important ones has negligible effect, validating the network's learned prioritization of angular information.

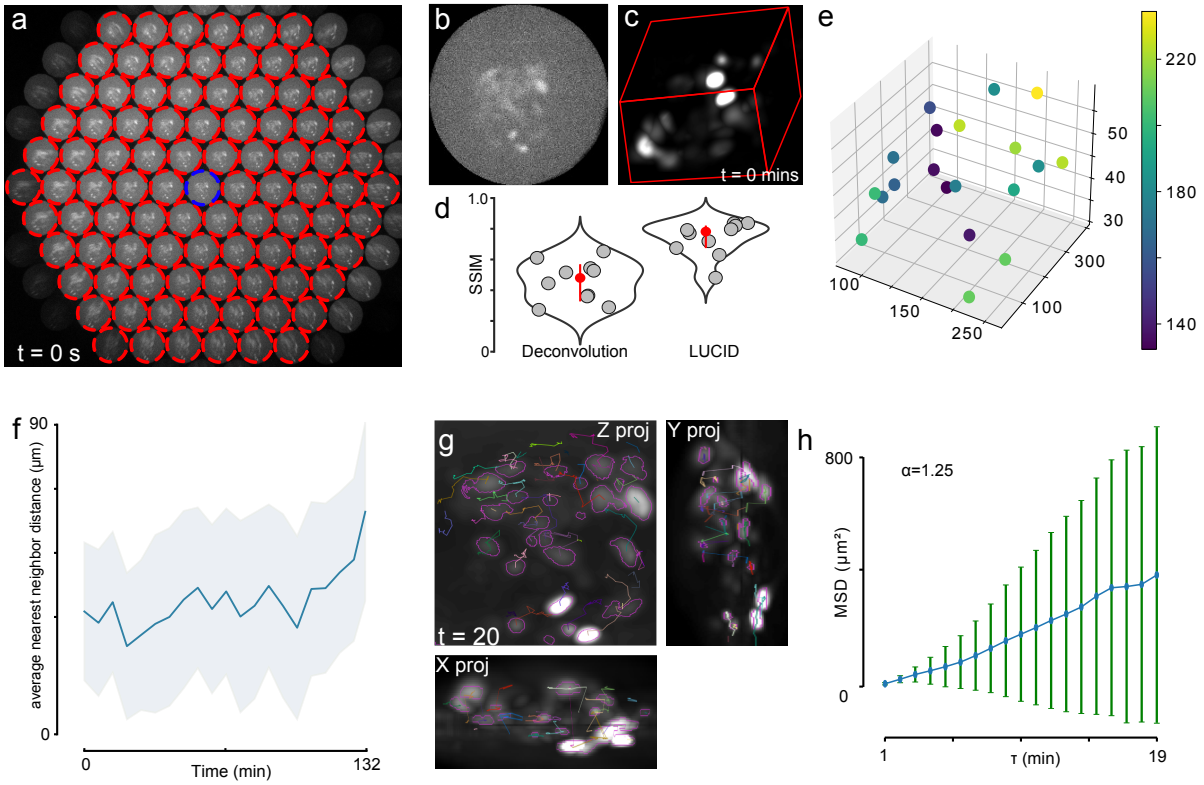
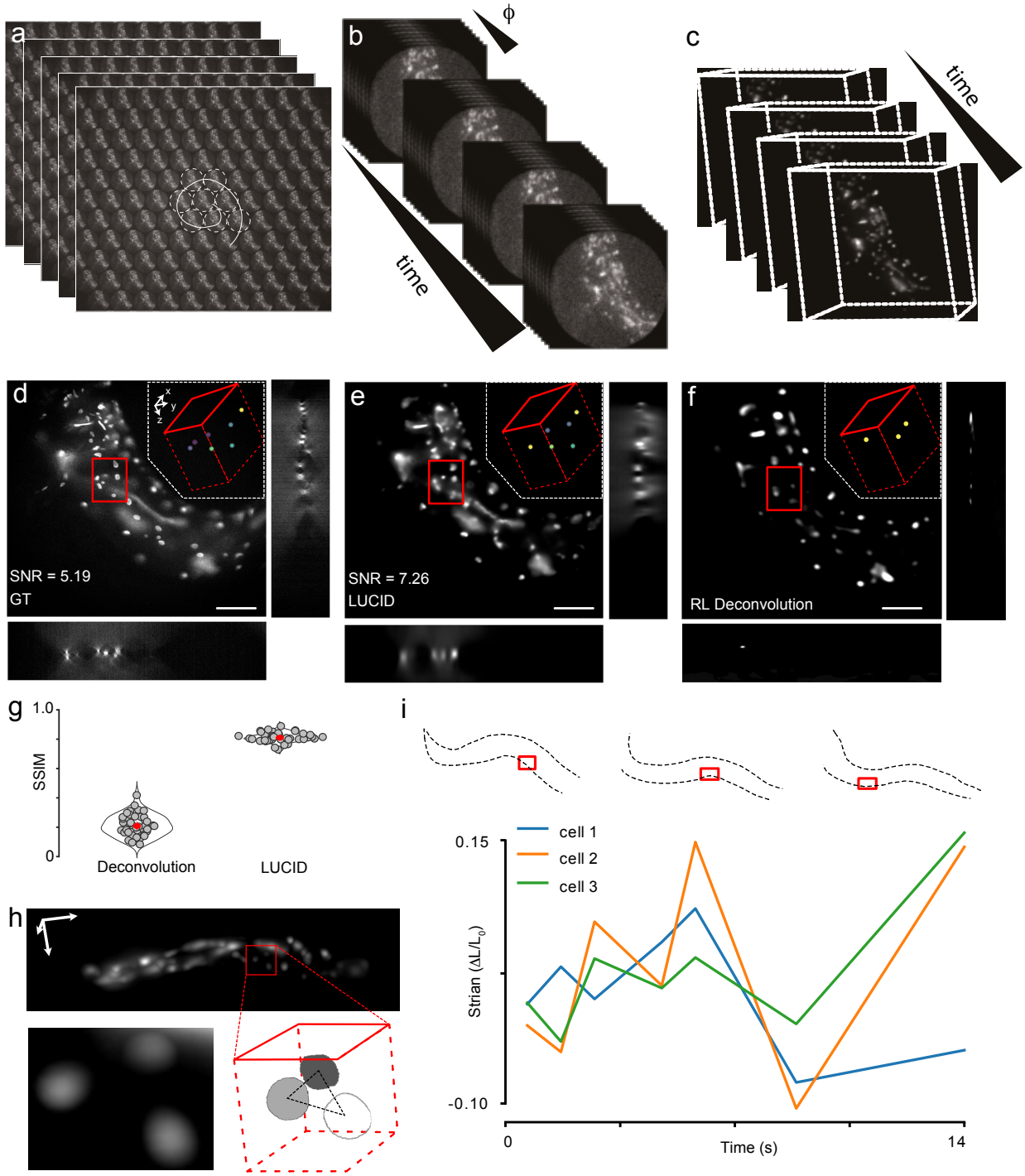


Figure 4: **LUCID enables 3D segmentation and tracking of mESC spheroids with nuclear labelling. a.** Raw FLF image, with all automatically detected perspectives in red outlines, and the central in blue. **b.** The central perspective (highlighted in blue in a) showing the raw intensity distribution at fMLA image plane. **c.** 3D volume of the reconstructed light field image at the initial time point (t = 0 min). See also Supplementary Video S. **d.** Distribution of the SSIM of the deconvolution and the LUCID reconstruction. **e, f,** 3D Representation of the proximity map of the segmented cells (e), with each dot indicating a segmented cell within the spheroid. (f) Average distance to each cell's three nearest neighbors as a during the time of the sequence shown in Supplementary Video . **g.** Maximum intensity projections of the reconstructed volume at t=0, overlayed with the trajectories of stem cells projected along the three principal axes (X, Y, and Z). **h.** Mean square displacement (MSD) of the tracked cells across the timelapse. The MSD curve follows a power-law behavior with an exponent $\alpha = 1.25$, indicating super-diffusive dynamics consistent with persistent, directed cell motion.

Figure 5 *(following page)*: **Nuclear dynamics in freely moving worms.** **a-c** Sketch illustrating the organization of EIs for inference. The algorithm first extracts the central EI and then retrieves the surrounding neighbors. This stack of EIs is processed by a trained neural network to perform real-time 3D reconstruction. The inference model enables real-time applications, allowing the reconstruction of dynamic processes over time. **d-f** Maximum intensity projection along z-direction of a *C. elegans* the fluorescently labeled GFP::LMN-1 (d) ground truth (GT), (e) the LUCID reconstruction and (f) the conventional Richardson-Lucy deconvolution for a single time point in Video . The inset shows the zoomed-in box displayed in red as a 3D segmented scene. Segmentation color indicates the average distance to the three nearest neighbors at the center of each cell. SNR, signal-noise ratio; scale bar = 20µm. **g** Structural similarity index (SSIM) comparing the reconstructions with our LUCID model and the RL deconvolution. **h** (i) Snapshot of a freely moving animal over the first 15 seconds of the timelapse. (ii) Close-up of three tracked cells along the body in 3D, highlighted in the red square of (i). **i** Mechanical tissue strain, plotted as the relative change in intercellular distance, $\frac{dL}{L_0}$, across time for the tracked cells to illustrate local body deformation during motion. Schematic on top indicates representative body posture during the timelapse.


a
b
ϕ
time
c
time
d
SNR = 5.19
GT
e
SNR = 7.26
LUCID
f
RL Deconvolution
g
1.0
SSIM
0
Deconvolution
LUCID
h
i
cell 1
cell 2
cell 3
0.15
Strian (ΔL/L0)
-0.10
0
Time (s)
14

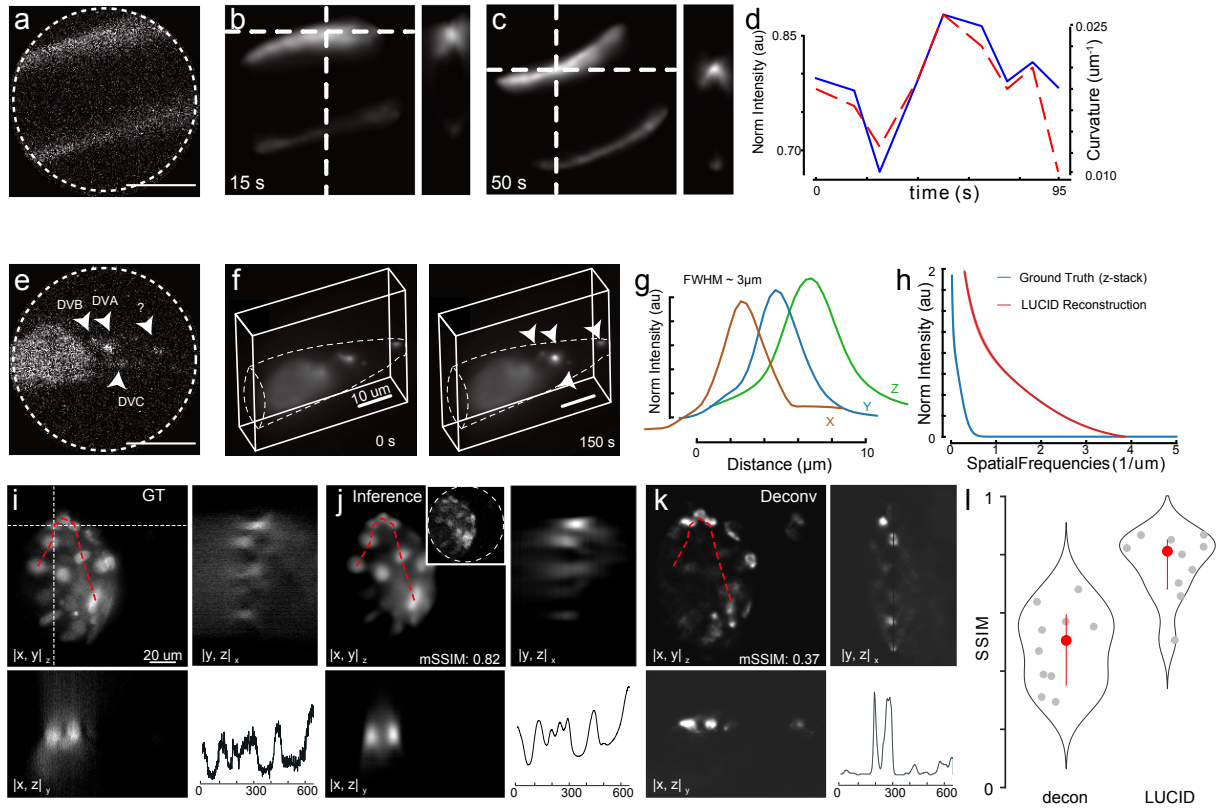


Figure 6: **Volumetric recordings of physiological activity with bioluminescent calcium indicators. a-d.** Calcium recordings in BWM of freely moving animals. (a) Central EI from the first frame in Video S of a freely-moving animal expressing a bioluminescent calcium-sensitive probe in the BWM. Scale bar = 50 µm. (b, c) Maximum intensity projection of the 3D reconstructions using LUCID, revealing high-calcium activity along the concave side during body bending at timepoints 15s (**b** and 50s **c**. (d) Correlation between bioluminescence intensity (blue) and curvature (dotted red) along the dorsal edge of the worm during locomotion. **e-g.** Neuronal calcium activity. (e) Central elemental image from the first frame in Supplementary Video S showing the dorsal tip of the tail with the dorsorectal ganglion containing DVA, DVB and DVC. Scale bar = 50µm. (f) Two frames (t=0 and t=150s) from Supplementary Video illustrate the isotropic reconstruction, with the DVA neuron clearly resolved and separated. Arrows indicate additional neurons in the tail region. (g.) Intensity profile of the DVA neuron in the x,y and z dimension. **h-i,** FLFM of mouse spheroids expressing H2B::mTurquoise2-TeNL to highlight their nuclei. Maximum intensity projections along z dimensions of the (i.) widefield z-stack acquired over 30 planes (GT, ground truth); (j.) the LUCID inference and (k.) deconvolution. In each case, the x,z and y,z planes are displayed along the dotted white lines shown in (i). Intensity distribution measured along the red line shown in the z-projection if panels i-k. **l.** SSIM comparing the restoration performed with the deconvolution and the LUCID network.

Materials and Methods

Instrument design

A custom-built optical microscope[5] was re-equipped with a transmitted LED light source (Thorlabs 590nm), a motorized stage and PiFOC for fast z-acquisition (Piezosystems Jena). Imaging was performed through a 40x 1.25 SiL immersion lens (Olympus # N4274100) and a single-photon resolving qCMOS camera (Hamamatsu Orca Quest). An epifluorescence path was added to enable multimodal acquisition of paired bioluminescent/fluorescent images for neuronal network training purposes using a Thorlabs 430nm LED. For Fourier light field imaging we redesigned the optical microscope introducing a dual optical branch to switch from a standard widefield microscope to Fourier light field imaging (Fig 2a). For this, we used a 1:1 relay lens with two AC508-150-A-ML lenses with a focal length of 150 mm to transpose the Fourier plane from the objective into the microlens array (MLA). The selected MLA for our system was the Okotech APH-Q-P1000-R14 with a focal length, pitch, size and radius of curvature of — mm, 1 mm, 14 mm × 14 mm × 1.5 mm and – mm respectively. We chose these specific properties to match the resolution we needed for our particular applications. Theoretically, this system, using classical reconstruction methods, is expected to achieve performances of 2.1 μm (R_{xy}), 16.5 μm (R_z), 88 μm (DoF), and 147 μm (FOV). To finish aligning the focal plane of the microlens array with the image plane, we used a relay lens composed of two lenses of 100 mm focal length AC508-100-A-ML. For the widefield light path, we used two right-angle mirror cage (Thorlabs, KCB2EC) with elliptical mirrors and a 200 mm tube lens to generate the image. To switch from one optical branch to the other one, redirecting the light, we employed two motorized flipper mirrors (Thorlabs, 8893-K).

Richardson Lucy Deconvolution for image reconstruction

Volumetric reconstruction of Fourier Light Field Microscopy (FLFM) data was performed using a model-based Richardson-Lucy (RL) deconvolution framework, adapted from the

computational methods previously described.[7,12] The reconstruction algorithm uses a forward model $\mathcal{H}$ representing the FLFM imaging system as a spatially varying convolution over axial depth, with each depth plane encoded by a 2D point spread function (PSF). The PSF stack H and its adjoint H^{T}, either experimentally measured or synthetically simulated based on the optical geometry, were preloaded into memory as depth-resolved kernel banks.

For a given raw lenslet image stack I, the forward model simulates the measured image by convolving each axial slice of an estimated object volume V with its corresponding PSF, and summing over depth. The inverse problem is solved iteratively using the RL update rule:

$$V_{k+1} = V_k \cdot \left(\frac{\mathcal{H}^{\mathrm{T}} \left(\frac{I}{\mathcal{H}(V_k)} \right)}{\mathcal{H}^{\mathrm{T}}(1)} \right), \tag{1}$$

where $\mathcal{H}$ and $\mathcal{H}^{\mathrm{T}}$ denote forward and backward projection operators, implemented through FFT-based convolutions on a GPU to accelerate computation. The object volume was initialized with a uniform guess and iteratively refined over 30 RL iterations. To minimize edge artifacts and reduce the data footprint, a centered sub-volume of fixed dimensions was cropped and saved every 10 iterations. The reconstruction yielded a 3D estimate of the object.

Attention neural networks for the reconstruction of Fourier light field images

We developed a deep learning-based pipeline for real-time volumetric bioluminescent imaging using an attention-based convolutional neural network (CNN) trained on synthetic, fluorescent and bioluminescent data. To generate the training dataset, we designed a dual-branch optical system that provides corresponding 3D wide-field images of immobilized samples, capturing their Fourier view perspectives using a micro-lens array (MLA). The images are acquired sequentially by switching between branches using motorized flip mirrors (Fig. 2a). For training, each perspective was used to extract features, which were then integrated into

subsequent convolutional layers. The network incorporates two attention mechanisms to preserve high-frequency information for enhanced reconstruction quality (Fig. 2b). First, a view-attention branch is added to each encoder block, addressing the loss of high-frequency content due to max-pooling layers. Second, attention gates are applied in the skip connections to combine spatial and angular information from the down-sampling and up-sampling paths. These soft-attention mechanisms reduce the activation of irrelevant input regions, minimizing redundant features.[24] The final layer of the network is configured to produce a specific number of slices that match the depth of the 3D target stack. We used a weighted loss function that combines two terms: pixel-wise mean squared error (MSE) and edge-aware error. The edge-aware term is computed by applying the Sobel operator to both the reconstructed and reference images and then taking the MSE between their edge maps. This encourages the network to preserve high-frequency structures, rather than simply weighting pixel errors near edges.[11] This combination drives the network to not only infer accurate intensity distributions but also generate sharper structures, resulting in higher resolution reconstructions.

Training procedure for Fourier light field data In the case of synthetic Fourier light field data, we first collected N different 3D stacks (of size 4096 px $\times$ 2304 px $\times$ 101 μm). The acquisition of the 3D stacks was made using an epifluorescence microscope with a 40$\times$/1.25-NA oil-immersion objective and a 430 nm excitation wavelength (Supplementary Figure 2a). We cropped the ROI from the acquired stack to account for the iris positioned in the image plane, which limits the FOV. As a result, all obstructed areas were removed to optimize computational memory usage. The same as before, we augment the synthetic data by performing linear transformations, e.g. flipping, rotating, and inverting the z axis to augment the dataset. We calculated the point spread function of the synthetic Fourier light field (PSF) based on wave optics theory by performing a forward projection onto the high resolution stacks[20] to generate the corresponding synthetic Fourier light field image to train

the network. The synthetic PSF was calculated to match our experimental setup; in other words, the synthetic PSF had the same amount of elemental images (EIs) as our experimental setup. The network training process follows a methodology similar to our earlier work,[5] where extracted views from the original light field image are transformed into a conventional 3D stack. We trained the network with a patience parameter of 10 epochs, which means that the training stopped if the loss function did not improve over 10 consecutive epochs. This training process required approximately 5 hours on a single graphical processing unit. As before, reconstruction performance was assessed using structural similarity (SSIM) and normalized root means squared error (NRMSE) metrics on unseen data.[5]

However, a network trained exclusively on synthetic data tends to produce artifacts or unintended patterns. To improve network performance, this preliminary model was expanded with experimental Fourier Light field data through transfer learning. This strategy improves performance because the synthetic datasets, while useful for providing large amounts of labeled data, do not fully capture the complexity, noise characteristics, and optical aberrations present in experimental acquisitions. By fine-tuning the model with real data, the network adapts its learned representations to the true data distribution, reducing the domain gap and enabling more accurate reconstructions.

The network training and data preparation consisted of the following steps:

- **Synthetic PSF generation:** A synthetic point spread function (PSF) was computed to match the experimental configuration.[20] Specifically, the number of elemental images (EIs) in the synthetic PSF matched the number in the experimental FLFM dataset.

- **Initial network training:** Following the approach of our previous work,[5] perspective views were extracted from raw Fourier light field (FLF) images and restructured into conventional 3D stacks. The model was trained with a patience parameter of 10 epochs, terminating training if the validation loss did not improve for 10 consecutive epochs.

- **Transfer learning with experimental data:** Since a model trained solely on syn-

thetic data introduced artifacts, we refined it using experimental data through transfer learning. We acquired a fluorescent image of *C. elegans* expressing a fluorescent protein under the control of the pan-neuronal *rab-3* promoter using the FLFM arm. The corresponding high-resolution 3D ground-truth stack was obtained with the widefield (WF) imaging arm on the same sample.

- **Perspective view extraction and calibration:** Raw FLFM images were processed to extract angular views. This required subpixel-accurate calibration of microlens positions, achieved by illuminating the sample with a homogeneous light source. A custom routine then automatically identified the lenslet array and extracted angular views from the raw image.

- **Preprocessing and augmentation:** The extracted perspective views were aligned and prepared as input–output pairs for supervised learning. We applied data normalization and augmentation to enhance generalization and reduce overfitting. The first stage of the network included a subpixel convolutional upsampling layer to enhance resolution while extracting relevant image features.

- **Widefield stack preparation:** WF images were cropped due to the pinhole located at the center of the relay lens, keeping only regions with relevant signal. The original WF volume (1300×1300×101 voxels) was downsampled to 800×800×101 to reduce computational burden without sacrificing structural integrity. This was further adjusted to 200×200×101 for compatibility with the model's transposed convolutional layers, allowing for a precise 4× up-sampling.

3D reconstruction with bioluminescence fourier light field data via transfer learning Volumetric reconstruction of bioluminescent Fourier light field (FLFM) images presents additional challenges compared to fluorescence imaging due to significantly lower photon counts and reduced signal-to-noise ratio (SNR). To address this, we generated an experi-

mental bioluminescent dataset using *C. elegans* expressing mTurquoise2 pan-neurally with the *rab-3* promoter, acquired at varying exposure times (1 s and 5 s). These FLF images—characterized by inherently lower SNR—were paired with corresponding widefield (WF) stacks serving as ground truth for supervised training.

Initial models pre-trained on fluorescence data were insufficient for bioluminescence, often producing blurry or noisy reconstructions that lack fine structural detail. To improve performance, we used a three-stage training strategy: (1) pre-training on synthetic data, (2) fine-tuning on fluorescence experimental data, and (3) final training on bioluminescence experimental data. This approach enabled the model to learn denoising and reconstruction jointly, resulting in markedly improved volumetric fidelity (Fig. 6).

In contrast, models trained solely on synthetic, fluorescence, or bioluminescence data performed poorly, frequently missing key features or amplifying noise (Supplementary Fig. 6). Traditional super-resolution Fourier-based methods[11] also failed to preserve depth integrity, often producing artificially flattened reconstructions and requiring longer processing times. Deep learning–based models, by comparison, yielded more accurate, denoised volumes and supported real-time reconstruction, highlighting the advantages of data-driven approaches for low-light volumetric imaging.

Sample preparation

Caenorhabditis elegans Animals were maintained using standard protocols[33] and grown at 20°C. Either larva at the fourth stage (L4) or young adults (YA) with the desired transgene were selected. For fluorescence imaging, worms were transferred to a drop of M9 buffer previously added on top of a 2% agarose pad.[34] For bioluminescence imaging worms were transferred to a drop of buffer containing 0.2 M citric acid, 0.4 M sodium phosphate dibasic, 2% dimethylsulfoxide (DMSO) and 0.1% Triton X-100 previously placed onto a 2% agarose pad. As substrate for the luciferase, fluorofurimazine (FFz)[35] 20 mM prepared in PBS was added to the worm suspension in a 1:1 proportion (10 mM final concentration). Transgenic

strains are described in Table 2.

Transgenesis. Molecular biology: All plasmids used for this study were constructed using the Gibson assembly method.[36] Generation of the plasmid expressing the TeNL in glutamathergic neurons and in BWM (pNMSB17 and pNMSB40 respectibvely) have been described before.[37] Plasmid for expression of the TeNL under the control of *rab-3* promoter (pNMSB133) was constructed by replacing the 15xUAS::delta *pes-10* promoter in pNMSB88 (Addgene plasmid #199321) with the *rab-3* promoter in pHW393 (Addgene plasmid #85583). Primers used for this assembly were 5'- TCCCCGGGATTGGCCAAAGG - 3' / 5'- CCATCGTTGTGAGTGATTTCAGCC - 3' for pNMSB88 and 5'- TGGCCAATCCCGGGGATCCTC - 3' / 5'- GGCTGAAATCACTCACAACGATGGATACG - 3' for pHW393.
Transgenesis: All transgenic animals used for this study were generated by microinjection[38] into the worm gonad with either of the DNAs previously specified at concentrations 30 ng/µl for pNMSB17, 5 ng/µl for pNMSB40 and 20 ng/µl for pNMSB133. pNMSB17 and pNMSB133 were coinjected with the IR87[39] as coinjection maker at 5 ng/µl. Plus DNA ladder 1kb from life Sciences up to a final DNA concentration of 100 ng/µl was used as DNA carrier in all cases. For bioluminescence imaging, the transgenic strains were crossed to incorporate the *bus-17(br2)* allele, facilitating the penetration of the substrate required for the reaction.[37] For the GFP::LMN-1 strain (MSB1136) a split GFP[40] was performed and the 11th barrel of GFP was introduced by means of the CRISPR/Cas9 technique at the amino end of the endogenous *lmn-1* gene. The experimental approach was using ribonuclases, repair templates with short homology arms (35bp) and the *dpy-10* system for identification of worms with edits.[41] Briefly, crRNA in the N-terminal region of *lmn-1* gene (5' - agaaaaATGTCATCTCGTAA - 3') was used as well as an ssODN containing left and right homology arms, a flexible linker and the GFP fragment (5' - tataattaactcttcagaaagcag cgagaaaaATGCGTGACCACATGGTCCTCCACGAGTACGTCAACGCCGCCGGAATC ACCGGTGGCGGCAAATTCTCATCTCGTAAAGGTACTCGTAGTTCTCGTATTGT

TACGCTAG - 3'). Both components, together with Cas9, tracrRNA were purchased from IDT Technologies. Mix components and concentration in the final mix were: 1.53 µM Cas9, 6.40 µM tracrRNA, 1.25 µM *dpy-10* crRNA, 5 µM *lmn-1* crRNA, 0.92 µM *dpy-10* repair ssODN and 2.20 µM *lmn-1* repair ssODN. The result strain was crossed to CF4588[40] to reconstitute the full GFP protein.

Mouse embryonic stem cells Stable, bioluminescent cell lines were established as described.[5] In short, a DNA coding for H2B::TNL (non calcium sensitive form of the TeNL) was used to integrate into mouse Embryonic Stem Cells. A CRISPR/Cas9 integration into the Rosa26 locus was attempted by the Tissue Engineering Unit at the Center for Genomic Regulation (CRG) in Barcelona. Briefly, CRISPR backbone PX330 was used to introduce the DNA for H2B::TeNL (non calcium sensitive form of the TeNL directed to the nucleus), 5'- ACTGGAGTTGCAGATCACGA - 3' was used as gRNA and mESC G4, F1 hybrid, superior developmental potential[42] as host cells. Full insert size was 4516 bp and right and left homology arms were 822 bp and 792 bp long respectively. Transfection was performed on 1 million G4 mESC using Lipofectamine™ 3000 Transfection Reagent (Thermo Fisher Scientific, ref. L3000001) mixed with 2.5 µg of total DNA, 0.625 µg of gRNA and 1.875 µg of circular HDR vector. Lipofection was carried out using a PenStrep-free media. Lipofectamine was removed after 24h. 5 days after lipofection, fluorescence was visible using an EVOS CFP cube. A first round of bulk FACS enrichment was performed. Nozzle 100 µm, To-Pro3 (Thermo Fisher Scientific, ref. T3605) as viability die. 1.37% of cells were positive for mTurqoise2. 13 days after lipofection, single cells were isolated using FACS and the same conditions as before. The 40% of cells positive for mTurqoise2 were sorted 5x96 plates, resulting in a total of 480 single clones. Screening revealed that 36 out of the 480 clones were expressing fluorescence but PCR testing with primers for the Rosa26 locus 5' - GGAAAAGTCTCCACCGGACG - 3' / 5' - TTGCATTCCAAAAGGAACCACC - 3' showed that integration had taken place randomly.

mESCs were maintained and expanded on 0.2% gelatin-coated dishes in mESC media composed of the following: Knock-Out DMEM (Thermo Fisher Scientific, 10829-018) supplemented with 15%Fetal Bovine Serum (FBS) (in-house mESC-tested), 1,000 U/ml Leukemia Inhibitory Factor (LIF, in-house generated), 1 mM Sodium Pyruvate (Thermo Fisher Scientific, 11360070), 1x MEM Non-Essential Amino Acids Solution (Thermo Fisher Scientific, 11140050), 50 U/ml penicillin/streptomycin (Thermo Fisher Scientific, 15140-122) and 0.1 mM 2-mercaptoethanol (Thermo Fisher Scientific, 31350010). Cells were cultured at 37°C with 5%CO2. Medium was changed every day and cells were passaged using 0.05% Trypsin-EDTA (Thermo Fisher Scientific, 25300054) and quenched 1:5 in DMEM supplemented with 10%FBS (Life Technologies, 10270106). Cells were tested monthly for mycoplasma contamination by PCR.

Spheroids were formed using ultra-low attachment (ULA) round-bottom 96-well plates were used to generate uniform spheroids. Cells were seeded at 5000 cells per well and allowed to aggregate for 24–48 h under standard culture conditions (37°C, 5% CO_2).

Spheroids were then collected and maintained in ES medium for imaging or downstream assays. For differentiation into embryoid bodies, LIF was removed, and spheroids were imaged in the lowlight microscope, in CO_2 independent medium and at 37°C using stagetop incubator (LCIBio).

Zebrafish Zebrafish (Danio rerio) were maintained as previously described.[43] Embryos were kept in E3 medium at between 25° and 31°C before experiments and staged according to morphological criteria[44] and hours post fertilization (hpf). 4hpf embryos were obtained from the transgenic background with a stable nuclear GFP expression (Tg(actb2:H2B-GFP)).[45] All protocols used have been approved by the Institutional Animal Care and Use Ethic Committee (PRBB–IACUEC) and implemented according to national and European regulations. All experiments were carried out in accordance with the principles of the 3Rs (replacement, reduction, and refinement).

Acknowledgement

The authors thank Aleksandra Pidde and Diego Ramallo for advice on programming, Gutavo Castro and Pablo Loza for help with the optical design and Montserrat Porta-de-la-Riva, Santiago Ortiz and Malak ElQuessny for help with biological sample preparation (C elegans, fish and mESCs, respectively). We thank Verena Ruprecht lab at the CRG Barcelona for sharing zebrafish embryos for Figure 1 d-f.

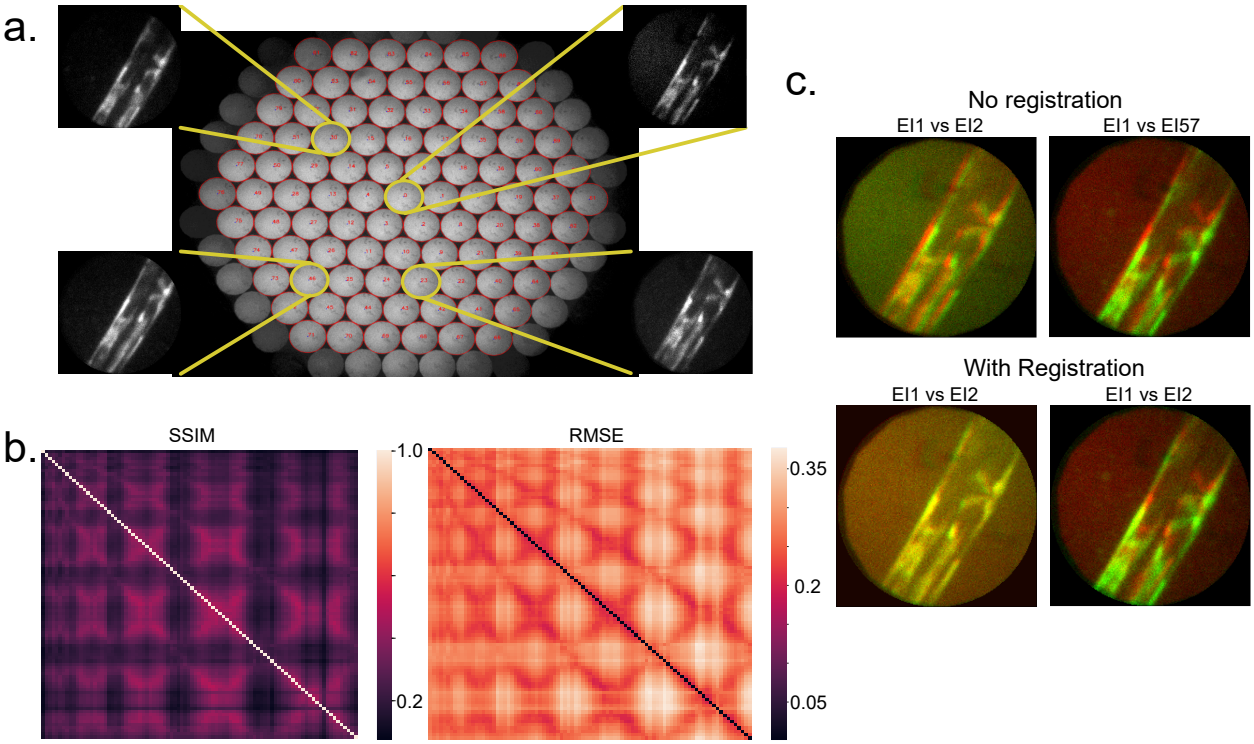
MK acknowledges financial support from the ERC (PoC "LowLiteScope", 101138041), Human Frontiers Science Program (RGP021/2023), MCIN/ AEI/10.13039/501100011033/ FEDER "A way to make Europe" (PID2024-157334OB-I00), "Severo Ochoa" program for Centres of Excellence in R&D (CEX2019-000910-S), from Fundació Privada Cellex, Fundació Mir-Puig, and from Generalitat de Catalunya through the CERCA and Research program and funding through H2020 Marie Skłodowska-Curie Actions (847517 to LFMC). ICFO is the recipient of the Severo Ochoa Award of Excellence of MINECO (Government of Spain). This work was supported in part by Oracle Cloud credits and related resources provided by Oracle.

Code availability All code is available at https://gitlab.icfo.net/NMSB/lucid

Supporting Information Available

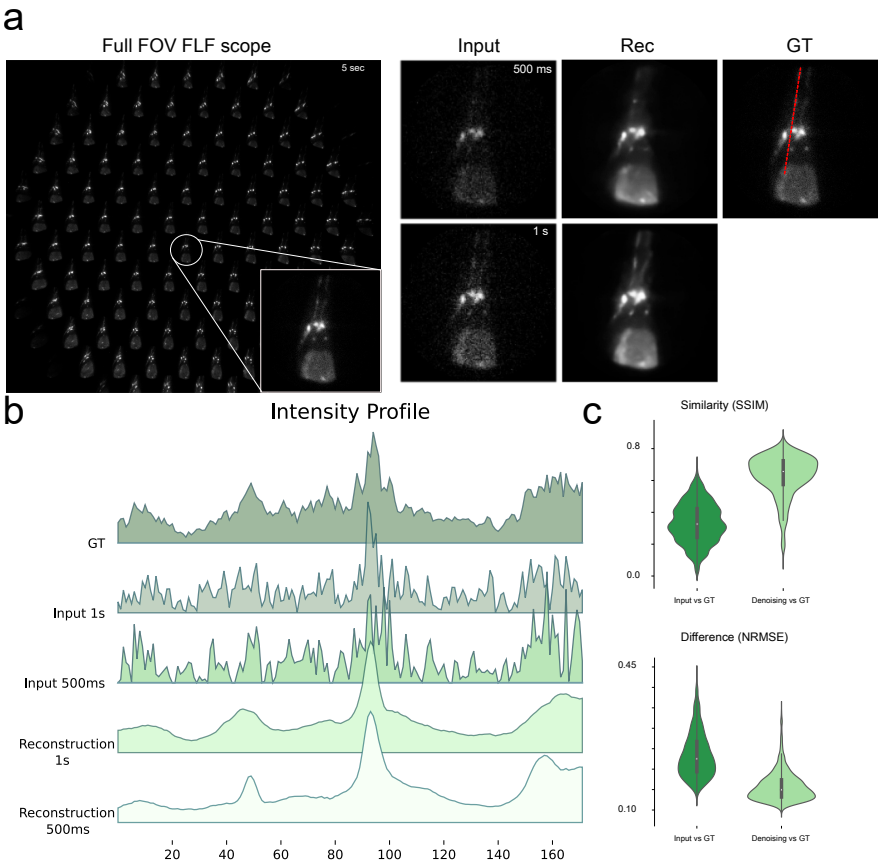
Supplementary Figures

Supplementary Figure S1



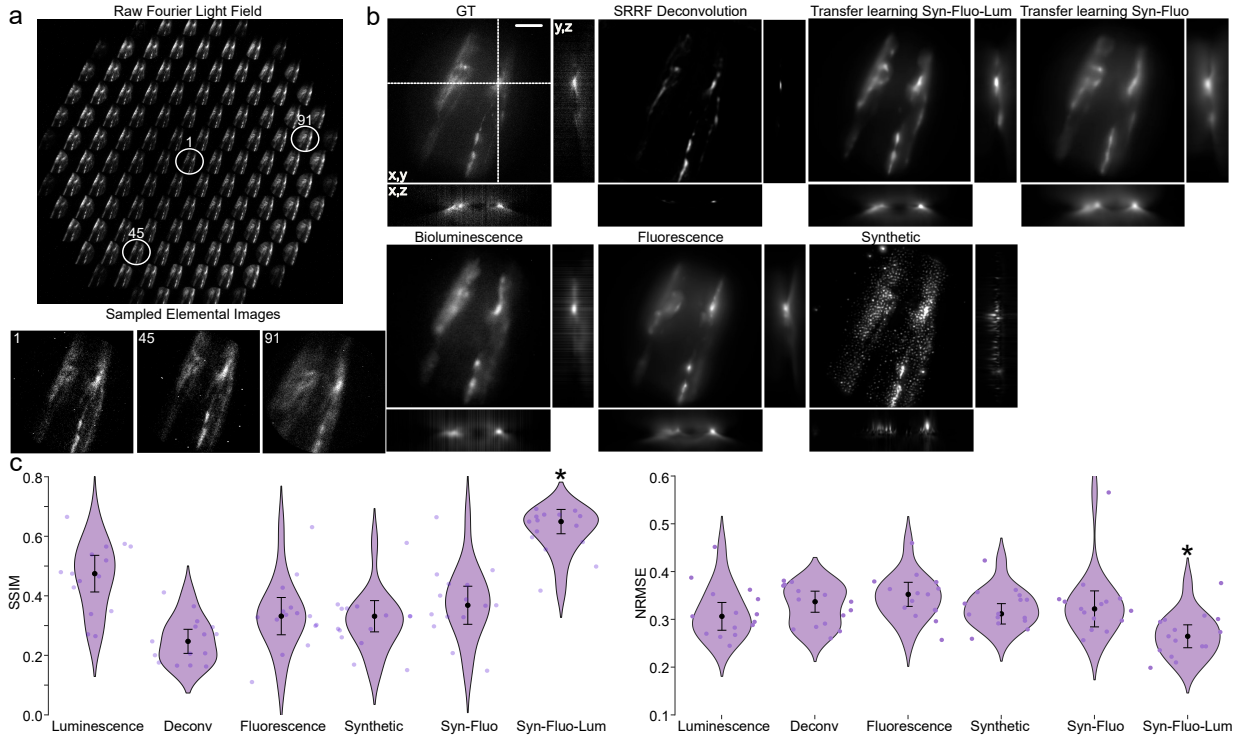
Supplementary Fig. 1: **Analysis of perspectives in FLFM. a**. Lenslet distribution showcasing four representative elemental images (numbered 0, 23, 30, and 46), each capturing unique information of *C. elegans* BWM. **b**. SSIM and RMSE heatmaps, highlighting low similarity across most perspectives except for self-comparisons along the main diagonal. **c**. Comparison of elemental images (BWM of *C. elegans*) after rigid registration, ensuring fair evaluation of information uniqueness.

Supplementary Figure S2



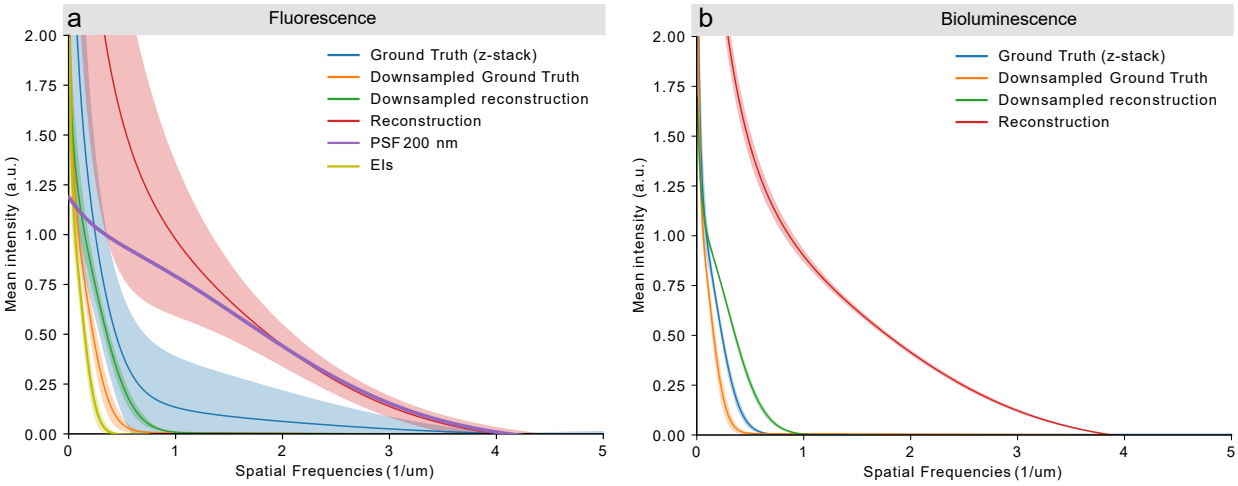
Supplementary Fig. 2: **Denoising of Fourier Light Field Elemental Images. a**. Experimental raw LF image with 91 perspectives of the head of a worm expressing the mTurquoise2 in glutamathergic neurons. Each of the images shows a different perspective from the sample. Noisy image input with exposure times of 500ms and 1s (top and bottom), then their respective restorations using a denoising model and an elemental image taken with a 5s exposure time. Along the head of the animal a red dotted line is shown **b** to see the intensity profile along the image. As seen the 2 inputs are very noisy before any posprocessing treament, howeven after the restoration is performe, a smooth signal distribution is shown, keeping the important regions of the image and removing useless information like noise in the background. **c**. The restoration achieves an approximately 2 times of improvement using the SSIM, NRMSE against the GT. N = 38.

Supplementary Figure S3



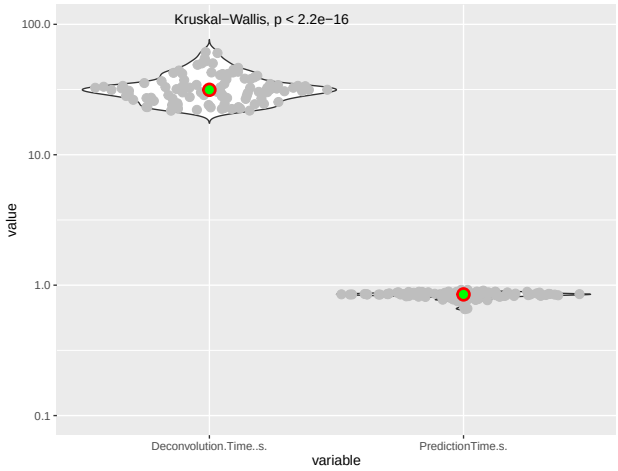
Supplementary Fig. 3: **Single-exposure Fourier light field microscopy of photon starved samples. a.** Raw Fourier light field image displaying the elemental images (EIs). Sampled EIs at positions (1), (45), and (91) are highlighted with white circles. **b.** Bioluminescent Fourier light field reconstructions obtained through different approaches, including super-resolution Fourier light field deconvolution via analytical methods, transfer learning at various training stages, and models trained with a single stage using bioluminescence, fluorescence, or synthetic datasets. **c.** Performance evaluation of individual models on unseen data using SSIM and NRMSE metrics (N = 15 images per model, no augmentation). The best-performing model is marked with an asterisk (*), showing a statistically significant improvement over the second-best model ($p < 5.78e\text{-}06$, paired t-test). Input image SNR = 3.029. Scale bar: 50 μm.

Supplementary Figure S4



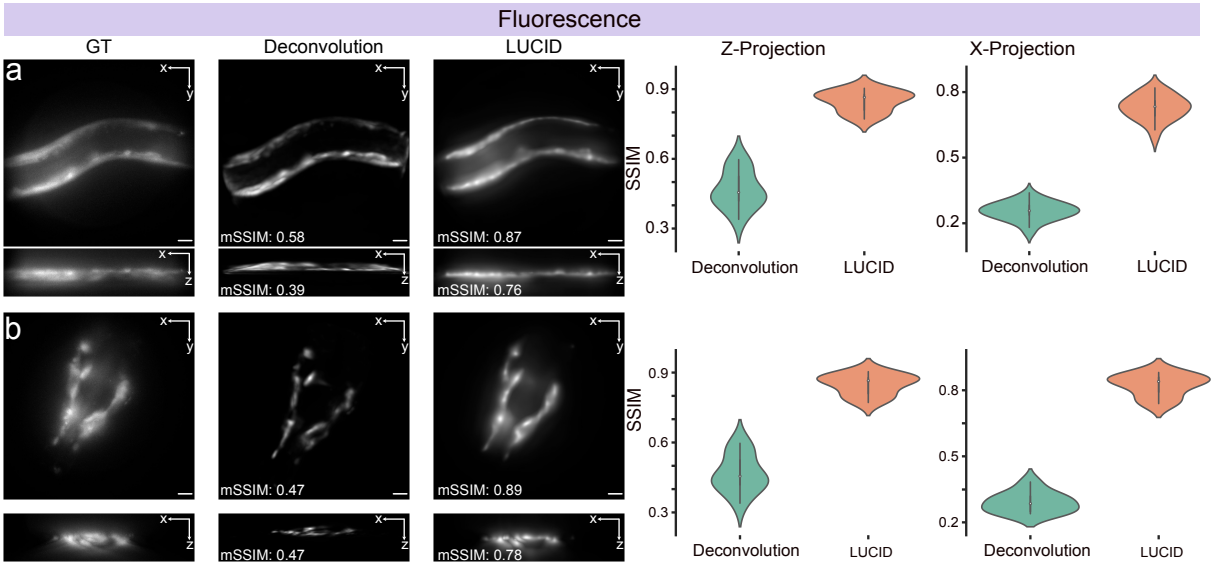
Supplementary Fig. 4: **Image resolution of the Fourier light field microscope. a.** Fluorescence: Azimuthal averages of the spatial frequencies from the FFT of the focal plane of fluorescence image stacks. These stacks correspond to either the original ground truth or a downsampled version to simulate reduced resolution. Two neural networks for Fourier light field reconstruction were trained: one using the downsampled ground truth (Downsampled reconstruction) and the other using the original ground truth z-stack (Reconstruction). Results are shown as mean mean±SD, with N = 10 technical replicates. **b.** Bioluminescence: Azimuthal averages of the spatial frequencies from the FFT of maximum intensity projections of bioluminescence image stacks, comparing the ground truth and the reconstructed light field stacks. The ground truth bioluminescence stack was acquired with a 10-second exposure time, which could introduce motion artifacts and reduced SNR, potentially degrading spatial resolution. Results are shown as mean±SD, with N = 10 technical replicates.

Supplementary Figure S5



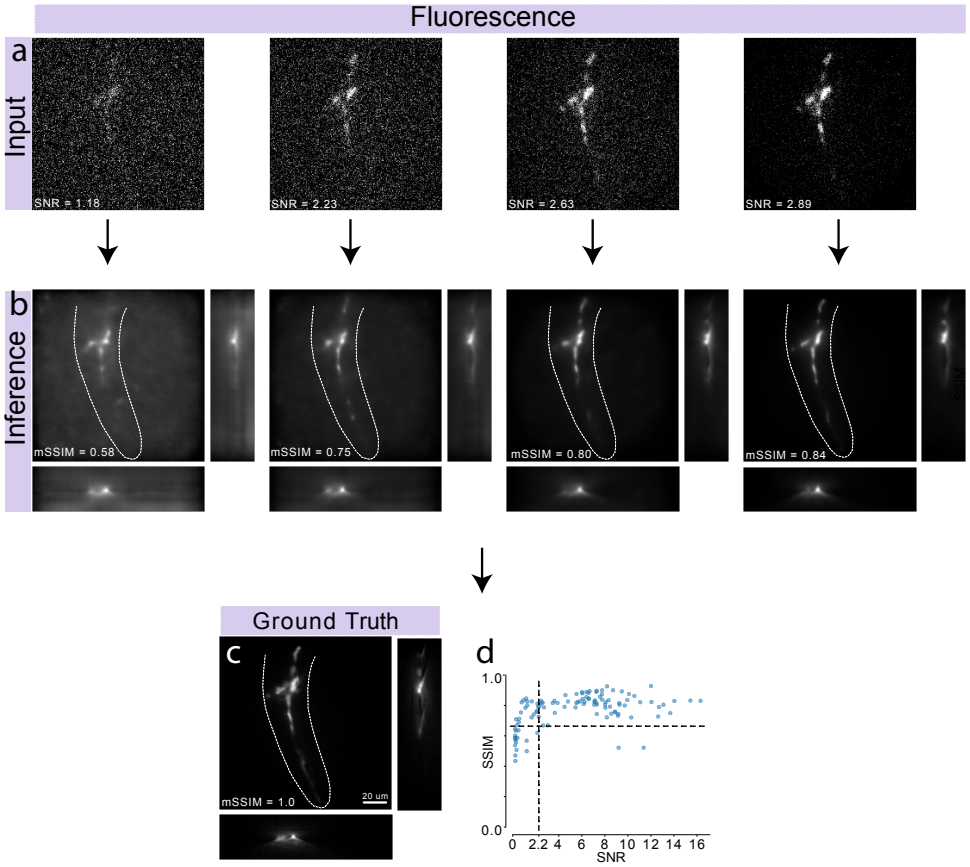
Supplementary Fig. 5: **Comparison of Execution Times: Deconvolution vs Neural Network** Box plot of the duration of the reconstruction of the FLF images using Richardson Lucy and our LUCID pipeline.

Supplementary Figure S6



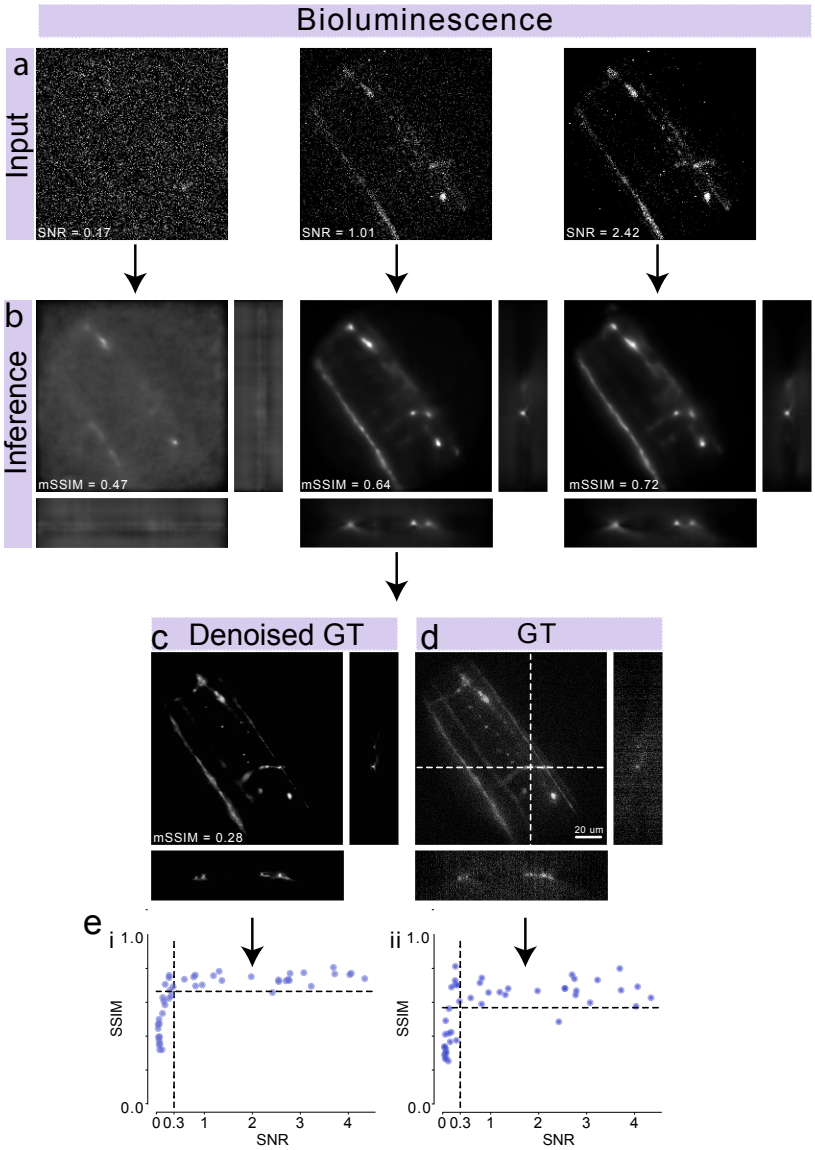
Supplementary Fig. 6: **Comparison of Neural Network and rad-FLFM Reconstructions of Fourier Light-Field Data. a.** Z- and X-projections of fluorescence Fourier light field reconstructions of the *C. elegans* body wall muscle. From left to right: GT, analytical reconstruction using rad-FLFM, our neural network approach using LUCID and the corresponding mSSIM comparison between the 2 algorithms. **b.** Z- and X-projections of the *C. elegans* neurons in the nerve ring, labeled with *rab-3p*::TeNL luciferase. From left to right: GT, rad-FLFM, and LUCID network reconstruction, with the associated mSSIM metrics for both projections. This comparison highlights the improved structural and volumetric fidelity achieved by our method.

Supplementary Figure S7



Supplementary Fig. 7: **Limits of restoration quality in FLF reconstruction. a.** Representative input images showing the central EI of fluorescent *C elegans* animal with neuronal labelling, acquired at different exposure time yielding varying signal/noise levels. **b.** LUCID reconstructions with their respective SSIM compared to the GT. Higher input SNR yields better structural similarity (SSIM) in the restored images. **c.** Ground Truth of the images studied in **a** and **b d.** SSIM plotted against input SNR, illustrating a threshold for effective reconstruction. Each dot represents a pair of images in the pipeline. GT = ground truth; SNR = signal-to-noise ratio. Scale bar = 20 μm.

Supplementary Figure S8



Supplementary Fig. 8: **Limits of restoration quality for bioluminescent FLFM reconstruction. a.** Input elemental images used for reconstruction, corresponding to different SNR levels, illustrating how input noise affects model performance. **b.** Orthogonal views of the FLF reconstructions, displayed alongside their respective mSSIM values. **c.** Orthogonal view of the denoised GT, obtained by applying a neural network-based denoising method to the raw GT, serving as a high-quality reference for comparison. **d.** Orthogonal view of the raw GT volume prior to denoising, showing the noise characteristics present in the original data. **e.** Plot of mean SSIM (mSSIM) versus input SNR, highlighting how reconstruction quality compares when using either the denoised or raw GT as a reference. GT = ground truth; SNR = signal-to-noise ratio. Scale bar = 20 μm.

Supplementary Tables

Supplementary Table 1 - Hyperparameters

Table 1: Summary of the different model configurations evaluated in the study.
Overview of the deep learning model configurations assessed in this study for fluorescence FLFM. Each entry includes the number of training images, presence or absence of data augmentation, indication of experimental status, use of transfer learning, network depth, whether any layers were frozen during training, application of denoising methods, and the specific network architecture employed.

Model	N images	Augmentation	Experimental	Transfer learning	Freezing	Denoising	Depth
1	748	Yes	Yes	No	No	No	2
2	748	N/A	Yes	No	No	No	3
3	748	N/A	Yes	No	No	No	4
4	748	N/A	Yes	No	No	No	2
5	1468	N/A	Yes	No	No	No	2
6	1468	N/A	Yes	No	No	No	4
7	1468	N/A	Yes	No	No	No	4
8	2188	N/A	Yes	No	No	No	4
9	2188	N/A	Yes	No	No	No	4
10	2188	N/A	Yes	No	No	No	4
11	2188	N/A	Yes	No	Yes	Yes	4
12	2188	N/A	Yes	Yes	Yes	Yes	4
13	2188	N/A	Yes	No	No	No	4
14	2188	N/A	Yes	Yes	No	No	4
15	2188	N/A	Yes	Yes	No	No	4
16	2188	N/A	Yes	Yes	No	No	4

Supplementary Table 2 - C elegans strains used in this study

Strain	Plasmid	Genotype	Expression	Figure
MSB577	pNMSB40	bus-17(br2)X, mirEx123[myo-3p::TeNL::unc-54 3'UTR]	BWM	2, 6, S1, S3, S6
MSB1082	pNMSB133 IR87	glo-4(ok623)V; mirEx462[rab-3p::NL250+rps-0p::hygroR]	pan-neuronal	2?, 3
MSB1309	pNMSB133 IR87	bus-17(br2)X; mirEx462[rab-3p::NL250+rps-0p::hygroR]	pan-neuronal	3, 6, S7, S8
MSB1136	-	mir83[GFP11::lmn-1])I; muIs253[eft-3p::sfGFP1-10::unc-54 3'UTR+Cbr-unc-119(+)],muIs252[eft-3p::wrmScarlet1-10::unc-543'UTR + Cbr-unc-119(+)]II;	Inner nuclear membrane	1, 2, 3, 5
MSB1217	pNMSB17 IR87	bus-17(br2)X; mirEx636[eat-4p::NL250+rps-0p::hygroR]	glutamatergic neurons	S2?, S7?

Supplementary videos

Supplementary Video S1

3D restoration of a short timelapse video of a developing zebrafish embryos at 50% epiboly. Zebrafish are transgenic for histone::GFP, and videos were deconvolved using Richardson-Lucy algorithm. Scale bar = 20 µm.

Supplementary Video S2

Short timelapse video of fluorescently labeled mouse embryo stem cells transgenic for H2B::NanoLantern.

Supplementary Video S3

Central elemental image of a freely behaving *Caenorhabditis elegans* animal, transgenic for H2B::nanolantern under fluorescence contrast.

Supplementary Video S4

Short timelapse video of freely behaving *Caenorhabditis elegans* transgenic for H2B::nanolantern under fluorescence contrast. Reconstructed using LUCID. Scalebar 20 µm.

Supplementary Video S5

3D restoration of a short timelapse video of freely behaving *Caenorhabditis elegans* transgenic for calcium sensitive Nanolantern in body wall muscles (*myo-3*p::TeNL) under bioluminescence contrast. Restoration was performed with the LUCID model. Scale bar = 20 µm.

Supplementary Video S6

3D restoration of a short timelapse video of freely behaving *Caenorhabditis elegans* transgenic for calcium sensitive Nanolantern under the control of the *rab-3* promotor in bioluminescence

contrast.

Supplementary Video S7

Central elemental image of a freely behaving *Caenorhabditis elegans* transgenic for H2B::nanolantern under fluorescence contrast. Scale bar 20 µm.

Supplementary Video S8

3D restoration of a timelapse video of freely behaving *Caenorhabditis elegans* transgenic for calcium sensitive Nanolantern under the control of the *rab-3* promotor in bioluminescence contrast, showing a section of the ventral nerve chord.

Bibliography

(1) Jacob, A. M.; Lindemann, A. F.; Wagenpfeil, J.; Geiger, S.; Layer, Y. C.; Salam, B.; Panahabadi, S.; Kurt, D.; Wintergerst, M. W.; Schildberg, F. A.; Kuetting, D.; Attenberger, U. I.; Abdullah, Z.; Böhner, A. M. Autofluorescence-based tissue characterization enhances clinical prospects of light-sheet-microscopy. *Scientific Reports* **2024**, *14*.

(2) Croce, A. C.; Bottiroli, G. Autofluorescence spectroscopy and imaging: A tool for biomedical research and diagnosis. *European Journal of Histochemistry* **2014**, *58*, 320–337.

(3) Laissue, P. P.; Alghamdi, R. A.; Tomancak, P.; Reynaud, E. G.; Shroff, H. Assessing phototoxicity in live fluorescence imaging. *Nature Methods* **2017**, *14*, 657–661.

(4) de-la Riva, M. P.; Morales-Curiel, L.-F.; Gonzalez, A.; Krieg, M. Bioluminescence as a functional tool for visualizing and controlling neuronal activity in vivo. *Neurophotonics* **2024**, *11*, 1–26.

(5) Morales-Curiel, L. F.; Gonzalez, A. C.; Castro-Olvera, G.; Lin, L. C. L.; El-Quessny, M.; de-la Riva, M. P.; Severino, J.; Morera, L. B.; Venturini, V.; Ruprecht, V.; Ramallo, D.; Loza-Alvarez, P.; Krieg, M. Volumetric imaging of fast cellular dynamics with deep learning enhanced bioluminescence microscopy. *Communications Biology* **2022**, *5*, 1330.

(6) Weigert, M. et al. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nature Methods* **2018**, *15*, 1090–1097.

(7) Prevedel, R.; Yoon, Y. G.; Hoffmann, M.; Pak, N.; Wetzstein, G.; Kato, S.; Schrödel, T.; Raskar, R.; Zimmer, M.; Boyden, E. S.; Vaziri, A. Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy. *Nature Methods* **2014**, *11*, 727–730.

(8) Wang, S.; Zhou, T.; Lu, Y.; Di, H. Detail-Preserving Transformer for Light Field Image Super-Resolution. *arXiv* **2022**, *arXiv:2201*.

(9) Cohen, N.; Yang, S.; Andalman, A.; Broxton, M.; Deisseroth, K.; Horowitz, M.; Levoy, M. Enhancing the performance of the light field microscope using wavefront coding. *Optics Express* **2014**, *22*, 727–730.

(10) Guo, C.; Liu, W.; Hua, X.; Li, H.; Jia, S. Fourier light-field microscopy. *Optics Express* **2019**, *27*, 25573–25594.

(11) Yi, C.; Zhu, L.; Sun, J.; Wang, Z.; Zhang, M.; Zhong, F.; Yan, L.; Tang, J.; Huang, L.; Zhang, Y. H.; Li, D.; Fei, P. Video-rate 3D imaging of living cells using Fourier view-channel-depth light field microscopy. *Communications Biology* **2023**, *6*.

(12) Stefanoiu, A.; Page, J.; Symvoulidis, P.; Westmeyer, G. G.; Lasser, T. Artifact-free deconvolution in light field microscopy. *Optics Express* **2019**, *27*, 31644.

(13) Wu, S.; Liu, W.; Hua, X.; Luo, J.; Wei, X.; Jia, S. PEARL: projection-estimation ac-

celerated Richardson-Lucy deconvolution for rapid volumetric imaging. *Optics Express* **2025**, *33*, 22685.

(14) Incardona, N.; Tolosa, A.; Saavedra, G.; Martinez-Corral, M.; Sanchez-Ortiga, E. Fast and robust wave optics-based reconstruction protocol for Fourier lightfield microscopy. *Optics and Lasers in Engineering* **2023**, *161*.

(15) von Chamier, L. et al. Democratising deep learning for microscopy with ZeroCostDL4Mic. *Nature Communications* **2021**, *12*, 1–18.

(16) Hallou, A.; Yevick, H. G.; Dumitrascu, B.; Uhlmann, V. Deep learning for bioimage analysis in developmental biology. 2021.

(17) Wang, H.; Rivenson, Y.; Jin, Y.; Wei, Z.; Gao, R.; Günaydın, H.; Bentolila, L. A.; Kural, C.; Ozcan, A. Deep learning enables cross-modality super-resolution in fluorescence microscopy. *Nature Methods* **2019**, *16*, 103–110.

(18) Schmidt, U.; Weigert, M.; Broaddus, C.; Myers, G. *Cell detection with star-convex polygons*; Springer International Publishing, 2018; Vol. 11071 LNCS; pp 265–273.

(19) Wu, Y.; Rivenson, Y.; Wang, H.; Luo, Y.; Ben-David, E.; Bentolila, L. A.; Pritz, C.; Ozcan, A. Three-dimensional virtual refocusing of fluorescence microscopy images using deep learning. *Nature Methods* **2019**, *16*, 1323–1331.

(20) Wang, Z.; Zhu, L.; Zhang, H.; Li, G.; Yi, C.; Li, Y.; Yang, Y.; Ding, Y.; Zhen, M.; Gao, S.; Hsiai, T. K.; Fei, P. Real-time volumetric reconstruction of biological dynamics with light-field microscopy and deep learning. *Nature Methods* **2021**, *18*, 551–556.

(21) Curiel, L. F. M.; Castro, G.; Gonzalez, A.; Lin, L.; Porta, M.; Ramallo, D.; Loza, P.; Krieg, M. Deep learning enhanced bioluminescence microscopy. Proc. SPIE 12204, Emerging Topics in Artificial Intelligence (ETAI). 2022; p 30.

(22) Lehtinen, J.; Munkberg, J.; Hasselgren, J.; Laine, S.; Karras, T.; Aittala, M.; Aila, T. Noise2Noise: Learning image restoration without clean data. *35th International Conference on Machine Learning, ICML 2018* **2018**, *7*, 4620–4631.

(23) Woo, S.; Park, J.; Lee, J.-Y.; Kweon, I. S. CBAM: Convolutional Block Attention Module. *arXiv* **2018**,

(24) Schlemper, J.; Oktay, O.; Schaap, M.; Heinrich, M.; Kainz, B.; Glocker, B.; Rueckert, D. Attention gated networks: Learning to leverage salient regions in medical images. *Medical Image Analysis* **2019**, *53*, 197–207.

(25) Zhang, Y.; Zhu, Y.; Liu, J.; Yu, W.; Jiang, C. An Interpretability Optimization Method for Deep Learning Networks Based on Grad-CAM. *IEEE Internet of Things Journal* **2025**, *12*, 3961–3970.

(26) Selvaraju, R. R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations From Deep Networks via Gradient-Based Localization. Proceedings of the IEEE International Conference on Computer Vision (ICCV). 2017.

(27) Suara, S.; Jha, A.; Sinha, P.; Sekh, A. A. Is Grad-CAM Explainable in Medical Images? Computer Vision and Image Processing. Cham, 2024; pp 124–135.

(28) Gong, J.; Yuan, Y.; Ward, A.; Kang, L.; Zhang, B.; Wu, Z.; Peng, J.; Feng, Z.; Liu, J.; Xu, X. Z. S. The C. elegans Taste Receptor Homolog LITE-1 Is a Photoreceptor. *Cell* **2016**, *167*, 1252–1263.e10, NULL.

(29) Ghosh, D. D.; Lee, D.; Jin, X.; Horvitz, H. R.; Nitabach, M. N. C. elegans discriminates colors to guide foraging. *Science* **2021**, *371*, 1059–1063.

(30) Ward, A.; Liu, J.; Feng, Z.; Xu, X. Z. S. Light-sensitive neurons and channels mediate phototaxis in C. elegans. *Nature Neuroscience* **2008**, *11*, 916–922.

(31) Das, R.; Lin, L. C.; Català-Castro, F.; Malaiwong, N.; Sanfeliu-Cerdán, N.; la Riva, M. P.-D.; Pidde, A.; Krieg, M. An asymmetric mechanical code ciphers curvature-dependent proprioceptor activity. *Science Advances* **2021**, *7*, 1–20.

(32) Hums, I.; Riedl, J.; Mende, F.; Kato, S.; Kaplan, H. S.; Latham, R.; Sonntag, M.; Traunmüller, L.; Zimmer, M. Regulation of two motor patterns enables the gradual adjustment of locomotion strategy in caenorhabditis elegans. *eLife* **2016**, *5*, 1–36.

(33) Stiernagle, T. Maintenance of C. elegans. *WormBook : the online review of C. elegans biology* **2006**, 1–11.

(34) de-la Riva, M. P.; Fontrodona, L.; Villanueva, A.; Cerón, J. Basic Caenorhabditis elegans methods: Synchronization and observation. *Journal of Visualized Experiments* **2012**, e4019.

(35) Su, Y. et al. Novel NanoLuc substrates enable bright two-population bioluminescence imaging in animals. *Nature Methods* **2020**, *17*, 852–860.

(36) Gibson, D. G.; Young, L.; Chuang, R. Y.; Venter, J. C.; Hutchison, C. A.; Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **2009**, *6*, 343–345.

(37) de-la Riva, M. P.; Gonzalez, A. C.; Sanfeliu-Cerdán, N.; Karimi, S.; Malaiwong, N.; Pidde, A.; Morales-Curiel, L. F.; Fernandez, P.; González-Bolívar, S.; Hurth, C.; Krieg, M. Neural engineering with photons as synaptic transmitters. *Nature Methods* **2023**, *20*, 761–769.

(38) Mello, C. C.; Kramer, J. M.; Stinchcomb, D.; Ambros, V. Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. *The EMBO journal* **1991**, *10*, 3959–3970.

(39) Radman, I.; Greiss, S.; Chin, J. W. Efficient and Rapid C. elegans Transgenesis by Bombardment and Hygromycin B Selection. *PLoS ONE* **2013**, *8*.

(40) Goudeau, J.; Sharp, C. S.; Paw, J.; Savy, L.; Leonetti, M. D.; York, A. G.; Updike, D. L.; Kenyon, C.; Ingaramo, M. Split-wrmScarlet and split-sfGFP: tools for faster, easier fluorescent labeling of endogenous proteins in Caenorhabditis elegans. *Genetics* **2021**, *217*.

(41) Paix, A.; Wang, Y.; Smith, H. E.; Lee, C.-Y. S.; Calidas, D.; Lu, T.; Smith, J.; Schmidt, H.; Krause, M. W.; Seydoux, G. Scalable and Versatile Genome Editing Using Linear DNAs with Microhomology to Cas9 Sites in Caenorhabditis elegans. *Genetics* **2014**, *198*, 1347–1356.

(42) George, S. H.; Gertsenstein, M.; Vintersten, K.; Korets-Smith, E.; Murphy, J.; Stevens, M. E.; Haigh, J. J.; Nagy, A. Developmental and adult phenotyping directly from mutant embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 4455–4460.

(43) Westerfield, M. The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (Danio rerio), 5th Edition. *University of Oregon Press, Eugene (Book)* **2007**,

(44) Kimmel, C. B.; Ballard, W. W.; Kimmel, S. R.; Ullmann, B.; Schilling, T. F. Stages of embryonic development of the zebrafish. *Dev Dyn* **1995**, *203*, 253–310.

(45) Wu, S. Y.; Zou, P.; Fuller, A. W.; Mishra, S.; Wang, Z.; Schey, K. L.; Mchaourab, H. S. Expression of cataract-linked γ-crystallin variants in zebrafish reveals a proteostasis network that senses protein stability. *Journal of Biological Chemistry* **2016**, *291*, 25387–25397.