TECHNIQUES AND RESOURCES

RESEARCH ARTICLE

DeepProjection: specific and robust projection of curved 2D tissue sheets from 3D microscopy using deep learning

Daniel Haertter[1,2,*], Xiaolei Wang[3], Stephanie M. Fogerson[4], Nitya Ramkumar[5], Janice M. Crawford[4], Kenneth D. Poss[4,5], Stefano Di Talia[5], Daniel P. Kiehart[4,5,*] and Christoph F. Schmidt[1,*]

ABSTRACT

The efficient extraction of image data from curved tissue sheets embedded in volumetric imaging data remains a serious and unsolved problem in quantitative studies of embryogenesis. Here, we present DeepProjection (DP), a trainable projection algorithm based on deep learning. This algorithm is trained on user-generated training data to locally classify 3D stack content, and to rapidly and robustly predict binary masks containing the target content, e.g. tissue boundaries, while masking highly fluorescent out-of-plane artifacts. A projection of the masked 3D stack then yields background-free 2D images with undistorted fluorescence intensity values. The binary masks can further be applied to other fluorescent channels or to extract local tissue curvature. DP is designed as a first processing step than can be followed, for example, by segmentation to track cell fate. We apply DP to follow the dynamic movements of 2D-tissue sheets during dorsal closure in *Drosophila* embryos and of the periderm layer in the elongating *Danio* embryo. DeepProjection is available as a fully documented Python package.

INTRODUCTION

Time-resolved three-dimensional fluorescence microscopy of transparent model organisms such as embryos of the fruit fly *Drosophila*, the nematode *Caenorhabditis elegans* or the zebrafish *Danio rerio* is a central tool in developmental biology. With modern techniques, the dynamics of entire organisms can be rapidly imaged as sequences of stacks of two-dimensional image slices, resulting in GBs of data per recording. Manual image processing is far too slow to mine such data, and potentially introduces bias. Convolutional neural networks, a form of machine learning, have been shown to far outperform conventional algorithms for visual feature extraction in many areas of research and engineering, including the life-sciences (Belthangady and Royer, 2019). Neural networks such as U-Net

[1]Department of Physics and Soft Matter Center, Duke University, Durham, NC 27708, USA. [2]Institute of Pharmacology and Toxicology, Göttingen University Medical Center, Göttingen 37075, Germany. [3]Advanced Light Imaging and Spectroscopy Facility, Department of Physics, Duke University, Durham, NC 27708, USA. [4]Department of Biology, Duke University, Durham, NC 27708, USA. [5]Department of Cell Biology, Duke University Medical Center, Durham, NC 27710, USA.

*Authors for correspondence (daniel.haertter@med.uni-goettingen.de, dkiehart@duke.edu, cfschmidt@phy.duke.edu)

D.H., 0000-0002-9582-6141; X.W., 0000-0003-2660-0859; N.R., 0000-0002-4086-4562; K.D.P., 0000-0002-6743-5709; S.D.T., 0000-0001-9758-7925; D.P.K., 0000-0002-5790-5623; C.F.S., 0000-0003-2864-6973

(Falk et al., 2019) can robustly segment and classify complex image features, defined by an initial training process.

We here present a new deep learning approach that allows us to automatically extract image content from dynamic curved 2D manifolds embedded in 3D image stacks of developing tissues. DeepProjection (DP) is an algorithm for structure-specific surface projections based on feature detection with convolutional neural networks. DP is a useful tool for developmental biology because, throughout phylogeny, cell sheet migration is a fundamental feature of morphogenesis in development and in wound healing. Cell sheets assume complex curved geometries and move to form internal organs and structures, including the neural tube, the gut and the heart in vertebrates.

To follow tissue sheet morphogenesis in living embryos, it is useful to extract selected sections of image content from different slices of each 3D stack to create 2D projections, while rejecting content from other planes before further processing. Past approaches have significant shortcomings. Maximum intensity projection (MIP) is simple and fast, but only works when fluorescence from the structures of interest dominates noise and off-target labels. MIP cannot differentiate subtle and low-intensity content of interest from bright content in other image planes. Manual omission of problematic slices or reduction of the overall imaging volume can also eliminate parts of the target content. Manual masking of individual slices is entirely impractical for long recordings. Other reported methods for *z*-projection use pixel value statistics, e.g. the minimum, the median or the sum of the pixels along the *z*-axis through the stacks, typically assuming that the structures of interest display the brightest fluorescence or have the sharpest contrast, but do not use information from neighboring pixels. The extracted 2D manifolds are easily distorted by single bright pixels, i.e. the masks are not continuous, have holes and no sharp edges. The extracted 2D manifolds are therefore not smooth. More sophisticated approaches can be classified into three categories: (1) smoothing of height maps derived by MIP (Blasse et al., 2017); (2) ranking of *z*-slices by visual pattern recognition of targeted tissue structures by edge filters (Erguvan et al., 2019), Fourier transforms or wavelet transforms (Forster et al., 2004); or (3) evaluating mean and variance of intensity distributions in the neighborhood of each pixel (Herbert et al., 2021). All these algorithms perform well only when the target structures are bright and clearly distinguishable from background noise. Bright fluorescent structures with a clear texture in the image background, such as auto-fluorescent yolk granules in *Drosophila* embryos, are not robustly discriminated against (Mavrakis et al., 2008). Most importantly, all these approaches are static, i.e. require manual parameter optimization for each new recording. Machine learning approaches, in contrast, can be trained to deal with a broad range of imaging conditions. One pioneering application of ML to developmental imaging is CSBDeep, a package for content-aware image restoration (Weigert et al., 2018). This package demonstrates

how convolutional neural networks can be used for combined projection and denoising of microscopy stacks of *Drosophila* wing development, by coupling a small convolutional network for 2D projection and a U-Net for subsequent denoising (Weigert et al., 2018; Falk et al., 2019). However, owing to the strong emphasis on denoising, original pixel intensities are not conserved in the output and the algorithm does not yield the manifold containing the tissue.
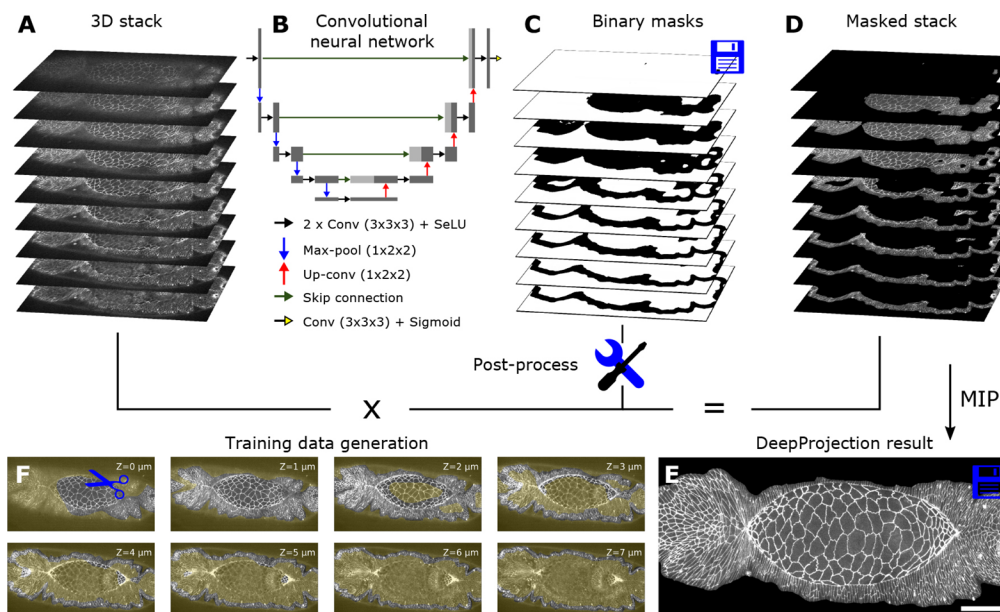
The key unaddressed challenge is to design an automated approach that can (1) detect defined fluorescent features that do not solely stand out by intensity or texture, and (2) project the entire 2D manifold in which the detected features reside without distorting intensities while completely rejecting content from regions outside of this 2D manifold. The crucial advantage of convolutional neural networks is that they can be trained to simultaneously detect various distinctive features of the target structure and then also select image content based on, but not limited to, the detected structures.

DP uses a convolutional neural network that analyzes features and textures in a 3D stack to create binary masks that contain only the regions specified as targets in the training data, in our case tissue layers containing sharp and distinctive cell boundaries. Thereby DP omits out-of-plane fluorescent structures and artifacts. We demonstrate DP using two models, dorsal closure in fruit fly (*Drosophila melanogaster*) development (Kiehart et al., 2000, 2017) and periderm development in zebrafish (*Danio rerio*) embryogenesis (Chang and Hwang, 2011; Eisenhoffer et al., 2017). DP predicts a single ZXY stack in just 1-10 s, depending on stack size. Significant gains in signal-to-noise ratio allow us to resolve even subtle structures. DP yields time-consistent results for time-lapsed movies as the algorithm detects persistent spatially extended features in the target manifold and is not deflected by temporary fluctuations and artefacts. DP also outputs masks that select 2D manifolds without modification of intensities in those planes. We show how these masks can be used to select image content from other fluorescent channels from the same 2D manifolds by extracting actin dynamics near the apical surface of amnioserosa cells during dorsal closure of *Drosophila*. We further show how the detected 3D geometry can be used to uniaxially flatten curved cell sheets.

RESULTS AND DISCUSSION

We developed DP to be broadly applicable to the rapid, automated processing of time-lapsed 3D recordings of developing embryos and tested it on recordings of dorsal closure in *Drosophila* embryos and periderm development in *Danio* embryos. At the core of DP is a custom-designed 3D convolutional neural network (CNN) that locally analyzes the image stack and classifies complex morphologies and textures (Fig. 1A,B, details in the Materials and Methods). For each application the user has to train a specific CNN model with a training dataset consisting of input stacks and manually generated ground truth. The CNN can simultaneously detect various spatially extended structures and robustly classifies each voxel into either target structure (e.g. a region with sharp cell boundaries) or background (e.g. fluorescent noise, artifacts, yolk granules, Fig. 1C). The resulting binary masks are then used to mask the input stack to exclusively contain the target structures (Fig. 1D). A maximum-intensity *z*-projection inside the predicted manifold then yields the final DP result (Fig. 1E). The predicted masks for each stack can be exported and subsequently applied to other fluorescent channels. Additionally, DP allows post-processing of binary masks to further tweak the algorithm result. DP was trained with pairs of image stacks and the corresponding manually created binary masks (Fig. 1F, details in the Materials and Methods).

In our applications to epithelial tissue sheets, only the regions of the stack containing tissue with sharp and distinctive cell boundaries are retained (Fig. 2A-H). Dorsal closure during *Drosophila* embryogenesis (Fig. 2A,C,E,F) occurs 12-15 h after egg laying (Sokolow et al., 2012; Kiehart et al., 2017). At this stage of development, the dorsal opening, which is left behind after germ band retraction, is covered by a curved sheet of squamous epithelial cells [amnioserosa (AS) cells]. The dorsal opening closes in ~3 h while AS cells subduct under the lateral epidermis (LE) or apoptose (Sokolow et al., 2012). We labeled cell-cell junctions with E-cadherin-GFP. Analyzing AS dynamics was complicated by the curved shape of the AS, highly auto-fluorescent yolk granules and subducted cells underneath the AS tissue (Fig. 2A,E). Tracking LE cells was impeded by low signal-to-noise ratios. DP clearly resolved AS and LE cell boundaries while yolk particles and gut cells were masked (Fig. 2C,E,F). During zebrafish development (Fig. 2B,D,G,H),
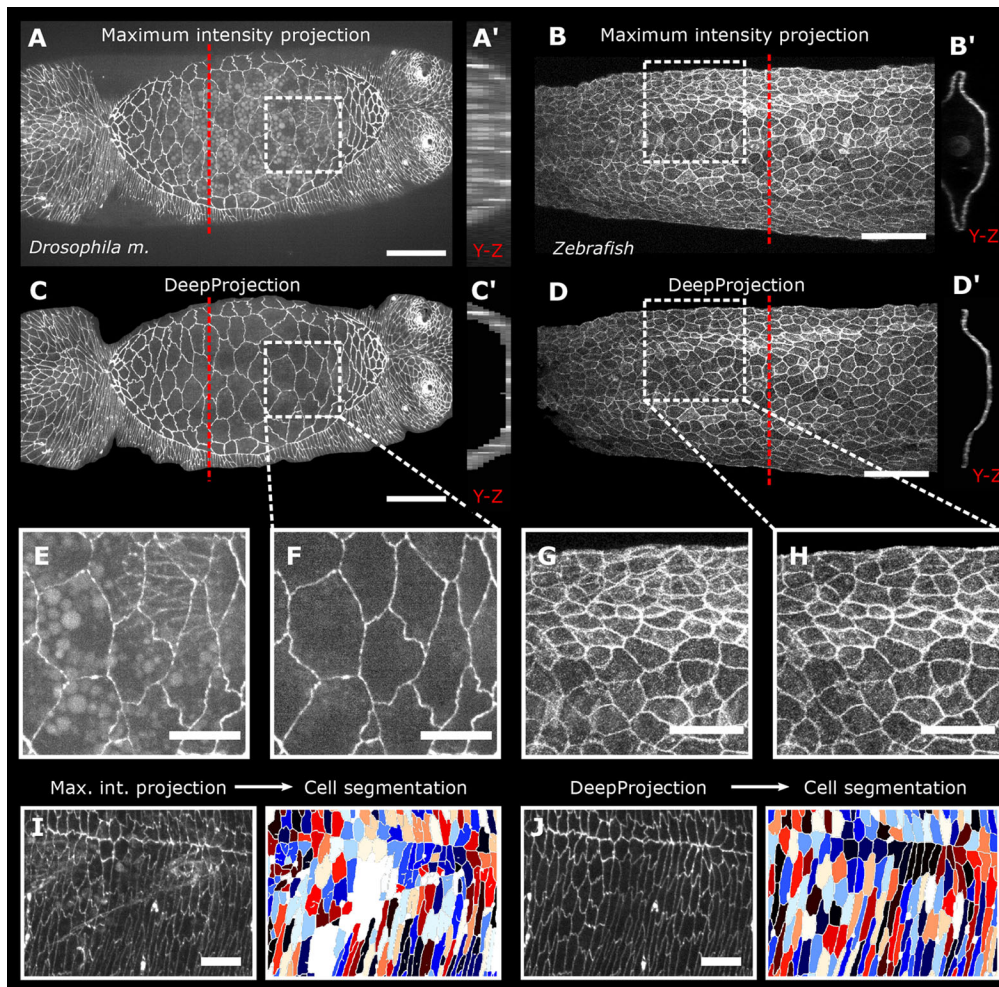

Fig. 1. DeepProjection algorithm. (A) Input 3D stack of *Drosophila* dorsal closure. (B) Convolutional neural network architecture. (C) Output of neural network showing binary masks (black=1, white=0). The predicted masks can be saved and applied to other fluorescent channels. (D) Multiplication of input stack with predicted binary masks yields a masked 3D stack containing only target tissue with sharp and distinctive cell boundaries. (E) MIP of masked 3D stack yields the result. (F) Illustration of manual training data generation. Yellow regions are cut out manually for each slide and only regions containing the target tissue are kept. Scale bar: 50 μm.

Fig. 2. Comparison of DeepProjection (DP) with maximum intensity projection (MIP). (A) MIP of a single stack (eight slices, 1 μm z-distance) of images of the dorsal opening of a *Drosophila* embryo during dorsal closure, cell boundaries labeled with Cadherin-GFP. (B) MIP of a single stack (53 slices, 2 μm z-distance) of images of zebrafish periderm labeled using *krt4*-directed lyn-EGFP fluorescence 1 day post fertilization (dpf). (A′,B′) *y-z* cuts of 3D image stacks at the red dashed lines in A and B. (C,D) DP results from the same stacks of *Drosophila* and zebrafish embryo. (C′,D′) The masked stack with the manifolds predicted by DP. (E,F) Enlargement of amnioserosal tissue in *Drosophila* comparing MIP (E) and DP (F), showing successful masking of yolk granules and gut tissue underneath the amnioserosal tissue. (G,H) Enlargement of a zebrafish embryo comparing MIP (G) and DP (H), showing masking of underlying epithelial tissue layer. (I,J) *Drosophila* lateral epidermis cell shape segmentation of MIP and DP results. Colors show cell labels. Scale bars: 50 μm in A,C; 100 μm in B,D; 10 μm in E,F; 50 μm in G,H; 15 μm in I,J.

the periderm covers the entire elongating embryo (Fig. 2B′). As the embryo narrows towards the posterior region, the left and right sides appear overlaid in the MIPs (Fig. 2B) when imaged laterally. This makes it difficult to distinguish the upper from the lower tissue layer. Furthermore, due to the strong curvature, projected cell shapes and areas are distorted. DP distinguished the tissue of interest from lower layers and the notochord, revealing the cell boundaries and making it possible to study cell shapes (Fig. 2D, G,H). As an example of subsequent processing, we segmented the strongly elongated cells of the *Drosophila* LE tissue with faint boundaries and achieved almost complete segmentation after application of DP, clearly superior to the application of a simple MIP, demonstrating the advantage of using DP prior to cell segmentation (Fig. 2I,J).

To demonstrate the capabilities of DP, we compared its performance with simple MIP, three published algorithms and manually generated ground truth: FastSME (FSME) (Basu et al., 2018), Local Z Projector (LZP) (Herbert et al., 2021) and CSBDeep (CSBD) (Weigert et al., 2018). We further tested Ilastik, a trainable pixel-segmentation algorithm not specifically designed for projections, as an alternative to our custom convolutional neural network (Berg et al., 2019). After multiple iterations of annotation and training, Ilastik was, unlike DP, not able to detect the target regions (Fig S1; supplementary Materials and Methods). We trained CSBDeep with the MIP of the masked stacks of the DP training data as ground truth (GT) using default parameters (200 epochs with

learning rate 4e-5). We then selected representative confocal stacks from recordings of *Drosophila* dorsal closure (*n*=8) and zebrafish periderm (*n*=4), which are distinct from the training data, and compared the respective results with manually created GT (Fig. 3A-C). The parameters of FSME and LZP were optimized individually for each stack.

For the dorsal-closure stacks, DP was able to reproduce the ground truth, while MIP and FSME failed to remove yolk granules and underlying gut tissue; LZP detected only parts of the AS tissue, and yolk granules leaked through; CSBD showed holes and yolk granules in the cell centers and the intensity gray values appeared distorted (Fig. 3A). The faint cell boundaries of the LE tissue were well detected by DP and LZP, while FSME did not perform better than MIP. CSBD showed opaque fog-like artifacts (Fig. 3B). Subducting cells along the seam under the LE were only discriminated against by CSBD and DP. CSDB distorted intensity values (Fig. 3B). For the zebrafish periderm, DP, LZP and CSBD were able to differentiate upper from lower tissue layers, in contrast to MIP and FSME (Fig. 3C). However, FSME, LZP and CSBD produced artifacts as black lines and cell boundary snippets, owing to the high local gradient of the tissue at the edge (Fig. 3C). CSBD again distorted intensity values, and high-frequency details inside the tissue manifold appeared smoothed or were missing (Fig. 3C). The CSBD algorithm attenuates or emphasizes certain features, which is evident in a plot of pixel grey values of the results against the GT (Fig. 3D). FSME and LZP yield pixel gray values
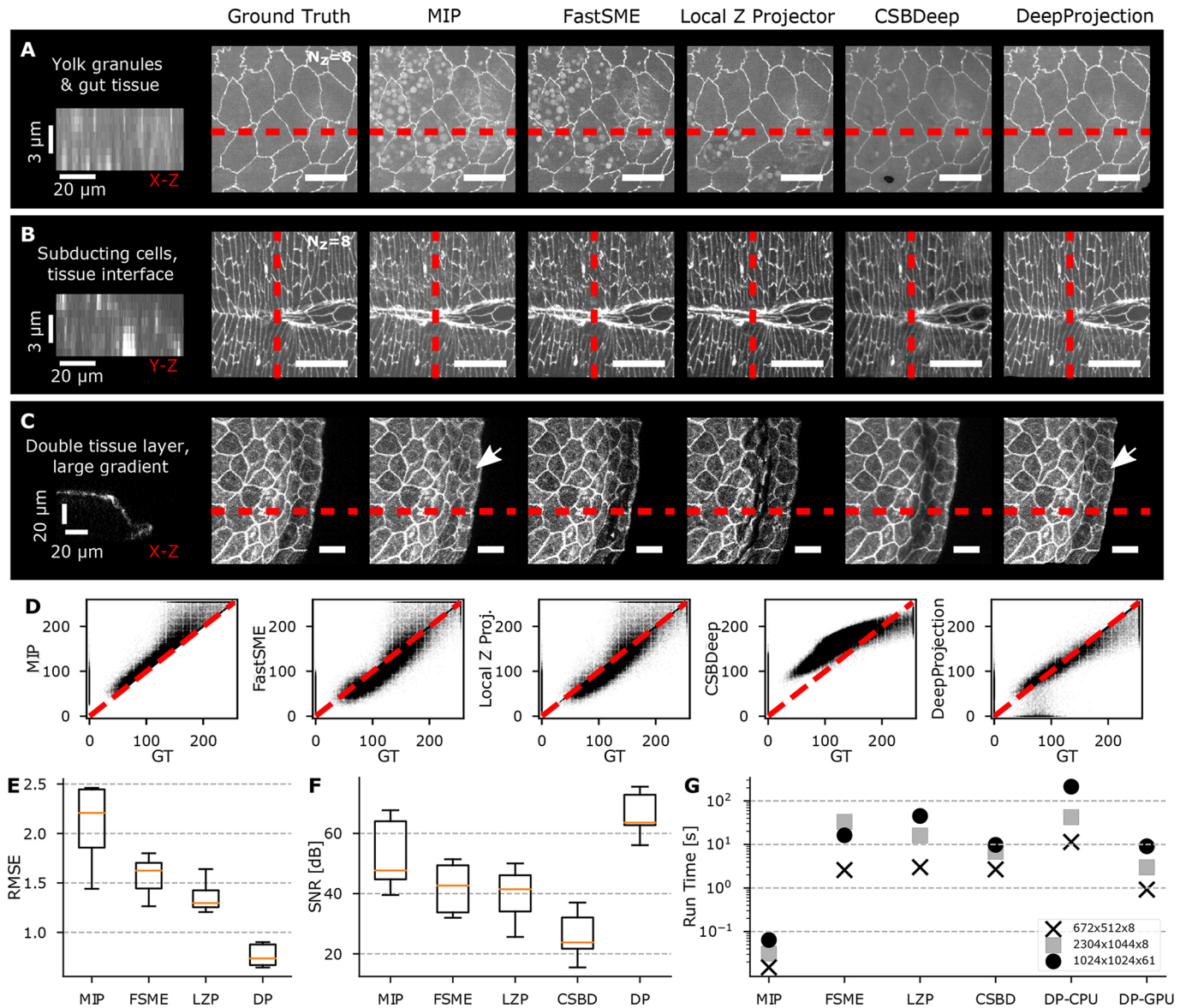
Fig. 3. Comparison of DP with published algorithms. (A) Comparison for a single confocal stack of amnioserosa (AS) tissue during early *Drosophila* dorsal closure, with auto-fluorescent yolk granules and gut cells. (B) Comparison for amnioserosa-lateral epidermis (LE) tissue interface and canthi during late *Drosophila* dorsal closure, with faint tissue cell boundaries, subducting cells and the interface of two different tissue types. (C) Comparison for zebrafish periderm (1 dpf), with large tissue gradient and second tissue layer underneath. Images on the left of A-C show vertical cuts at the red dashed lines, each averaged over 10 pixels perpendicular to the line. (D) Pixel-wise intensity scatter plot of algorithm results against ground truth to check for distortion of fluorescent gray value. (E) Root-mean-square errors of algorithm results relative to ground truth. (F) Signal-to-noise ratio of algorithm results with ground truth as reference. (G) Log-scale plot of algorithm run time for three different stack sizes. In box-and-whisker plots, boxes show the median (red) with IQR, with whiskers extending to the 5th and 95th percentiles. Scale bars: 20 µm.

scattering around the GT, while DP results were almost identical to GT. CSBD, however, nonlinearly distorts the pixel values, which makes it impossible to quantitate protein concentrations (Fig. 3D).

We further evaluated the performance of all methods in three ways.

(1) We calculated the normalized root-mean-square errors (RMSE) of the z-height maps $Z(x, y)$ with respect to the ground truth:

$$RMSE = \sqrt{\frac{\sum_{x,y}\left(Z_R(x,y) - Z_{GT}(x,y)\right)^2}{W \cdot H}}.$$

As DP yields masks with potentially more than one slice selected per x-y pixel, we created unique z-height maps for DP and ground truth by selecting the z-height corresponding to the maximum intensity inside the manifold. The CSBD package does not output a z-map or binary masks. DP strongly outperformed all other algorithms (Fig. 3E).

(2) We calculated the signal-to-noise ratio (SNR) of the reconstruction results I_R with respect to the ground truth I_{GT}:

$$SNR = 20 \log\left(\frac{\sigma^2(I_R)}{\sigma^2(I_R - I_{GT})}\right).$$

DP performed significantly better than both FSME and LZP (Fig. 3F). Interestingly, even MIP performed better than FSME and LZP, as all bright features, desired and undesired ones, are conserved by MIP. FSME and LZP both create a smooth z-map with only one selected plane per x-y pixel, whereas DP predicts a set of binary masks which embed the tissue of interest. When more than one slice is selected for a given x-y pixel, DP then uses MIP inside this embedding to obtain the final 2D projection result. When the structures of interest (in this case fluorescent cell boundaries) span multiple z-slices, MIP and foremost DP yield better SNR results by including relevant signal from more than one slice. Owing to the non-linear distortion of gray values, CSBD scored poorly (Fig. 3D).

(3) We assessed the run time of algorithms on exemplary stacks of three different sizes (Fig. 3G). MIP was very fast due its simplicity. DP required about 1 s to predict a stack with dimension 8×640×512 pixels, when run on a graphics card (GPU), three times faster than FSME and LZP. When run on the CPU, DP prediction took around ten times longer than on the graphics card but was still within a practical range. CSDB run times on a GPU were slightly longer than DP.

To demonstrate the option of mask transfer to other simultaneously recorded fluorescent channels, we performed dual-color imaging of dorsal closure labeled for E-cadherin and filamentous actin. The MIP of the actin channel shows intracellular actin networks and actin-rich filopodial cores (Fig. 4A). However, it is not possible to distinguish
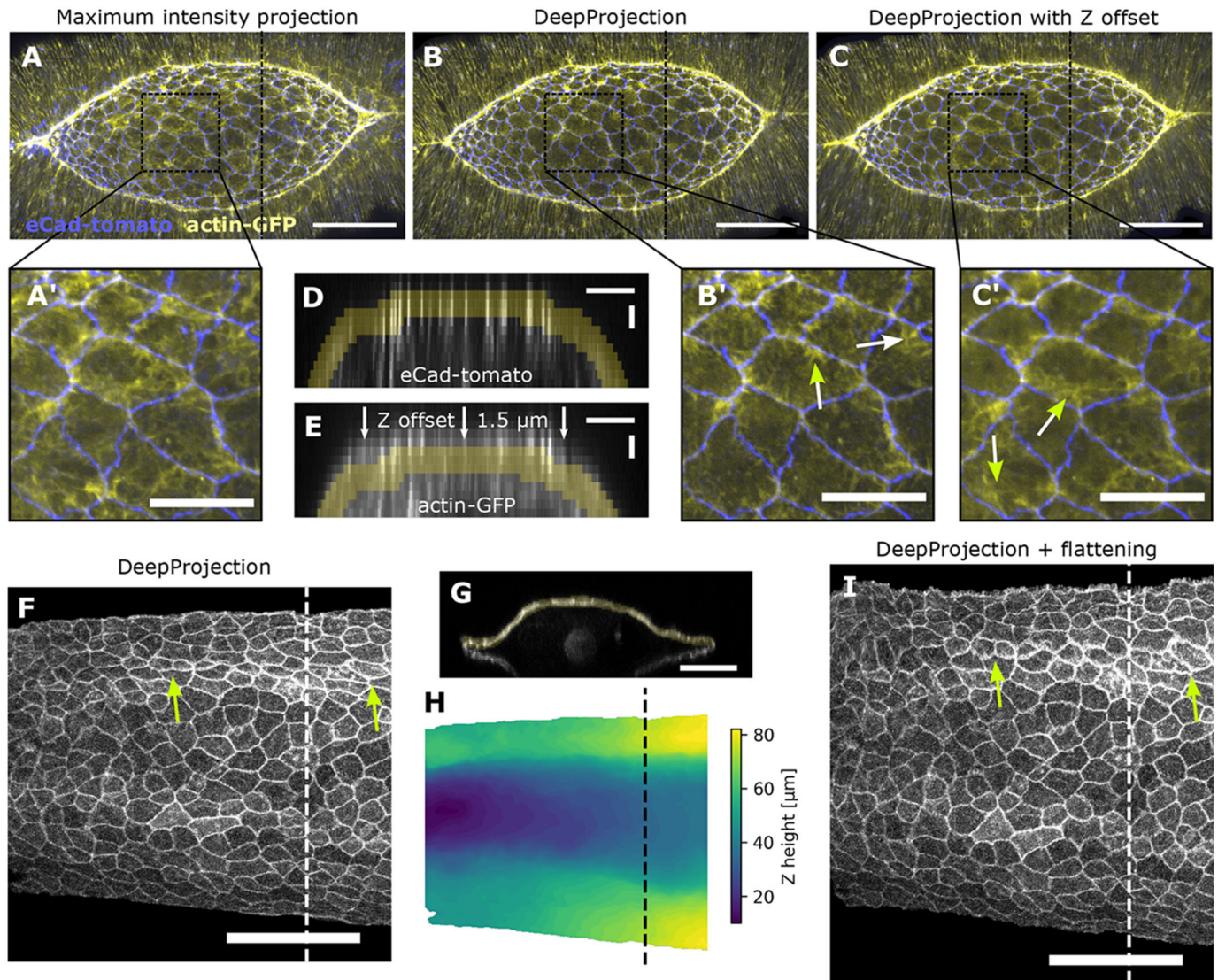

Fig. 4. Mask transfer to project content from other fluorescent channels and tissue flattening. (A) MIP of dual-color confocal stack of *Drosophila* embryo during mid-stage of dorsal closure with labeling via E-cadherin-tomato and actin-GFP. (B) Result of applying the masks, predicted by DP from E-cadherin cell boundaries to the actin channel, showing mainly actin-rich filopodia at the apical surface of cells. (C) Applying the same masks with a z-offset of 4 pixels/2 μm to the actin channel, showing contractile actin networks underneath the apical surface. (A′,B′,C′) Enlargements of A-C. Arrows indicate filopodia (B′) and contractile actomyosin network (C′). (D) y-z cut of E-cadherin-tomato channel (at dashed line in A-C) with binary mask predicted by DP. (E) y-z cut of actin channel with binary masks with z-offset. Arrows indicate cells at region with large curvature with distorted cell shape in projection. (F) DP result of zebrafish periderm development (1 dpf). (G) x-z cut along white dashed line in F. The manifold predicted by DP is highlighted in yellow. (H) z-height map calculated by averaging the positive indices of the manifolds at each x-y position. The z-map was subsequently smoothed with a mean filter with kernel size 5×5 pixels. (I) Result of flattening algorithm applied in y direction. As highlighted, the flattening reveals the true shape and area of cells at positions with large gradient. Arrows indicate cells with corrected shape compared with F due to flattening. Scale bars: 50 μm in A-C,G; 20 μm in A′-C′; 20 μm in y, 1.5 μm in z in D,E; 50 μm in F,I.

between apical actin, which is responsible for the contraction of apical cell areas (Ma et al., 2009; Duque and Gorfinkiel, 2016), and actin elsewhere in the cells. We next used DP to predict binary masks using the cell boundary information captured in the E-cadherin channel and then applied the masks to the actin channel (Fig. 4B,D). As highlighted in Fig. 4B′, the predicted masks include the apical surface with actin networks and filopodia. A comparison between the DP and MIP results shows additional pronounced actin structures at the basal surface of cells visible in MIP, but not in DP (Fig. 4A′,B′). In order to extract the basal actin, we shifted the binary masks by four pixels (corresponding to 2 μm in the z-direction) (Fig. 4E), and then applied them to the actin channel (Fig. 4C). DP thus allows us to not only extract image content from other channels from the predicted 2D manifolds, but also from other parallel planes, offset in the z-direction.

If a curved cell sheet displays steep gradients, cell shapes and areas are distorted in z-projections (Fig. 4F,G). By processing the binary masks, DP extracts the local z-height of the target tissue (Fig. 4H, Movie 2). Based on the z-height map, we implemented an flattening algorithm that successfully straightens the regions of the zebrafish embryo with high gradients in x, revealing true cell sizes and shapes (Fig. 4I). Our approach is, in this respect, similar to previously reported unrolling and cartographic approaches (Heemskerk and Streichan, 2015).

So far, we have focused on the projection of individual confocal stacks. When evaluating time-lapse recordings of developing embryos, time consistency of the projection method becomes important, i.e. the differences between consecutive frames reflect only real morphological changes and do not show projection artifacts as flickering or missing values. We applied DP to time-lapse recordings of dorsal closure (Movie 1). Even though the stacks were predicted one-by-one, and no information was propagated between consecutive time points, the results were time consistent, demonstrated by the stable position of tissue edges (Movie 1). DP robustly detects time persistent image features, of both low and high spatial frequencies and is thus not perturbed by time-varying noise or decreasing fluorescent intensity due to photobleaching. To further improve time consistency and to eliminate flickering artifacts that occasionally occurred, we added the option to average the binary masks using rolling-window mean filtering of voxels over multiple frames. Consecutively, the filtered masks are again binarized and used for projection, yielding improved projection results (Movie 1).

Conclusions

DP uses a convolutional neural network to selectively extract and project image content from curved 2D manifolds embedded in 3D confocal stacks. DP detects complex features that are in each application specified by user-annotated training, which makes it possible to mask even highly fluorescent artifacts while faithfully detecting weakly fluorescent structures of interest. DP could extract and project image content from dynamic curved tissue sheets in both *Drosophila* and zebrafish embryos while masking background content and noise. Image processing with DP greatly simplifies the segmentation and tracking of individual cells in subsequent processing steps. Original fluorescence intensity values in the selected manifold are strictly preserved for quantitative analysis. DP significantly outperformed the alternative algorithms we tested and created time-consistent results for time-lapsed recordings. DP is substantially faster than most published algorithms. Owing to its universal architecture, DP is not limited to the analysis of epithelial tissues but can be applied to extract any 2D manifold from 3D data

when properly trained for the features of the target manifold. For high-content imaging pipelines, DP can rapidly and robustly compress data from stacks to single images to save storage space without losing information from the target manifolds. Deep learning algorithms such as DeepProjection constitute a major leap in the capability to process large amounts of imaging data and will enable researchers to rapidly mine data and rigorously quantify complex morphogenetic processes.

MATERIALS AND METHODS
Preparation and imaging of *Drosophila* embryos
Cell junctions were labeled with either DE-cadherin-GFP or DE-cadherin-mTomato (labeling *Drosophila* E-cadherin, which is concentrated in cell-cell junctions), both knock-in lines under control of the endogenous promoter (Huang et al., 2009). F-actin was labeled with the GFP-moesin actin-binding domain, expressed under the control of the *spaghetti squash* promoter in the sGMCA line (Kiehart et al., 2000). All stocks were maintained at room temperature or 25°C on standard cornmeal/molasses fly food or in embryo collection cages with a grape juice agar plate and yeast paste. Embryos were collected either 2-4 h after egg lays and aged overnight at 16°C or from overnight egg lays at 25°C. To remove the chorion, embryos were incubated in a 50% bleach solution for 1.25 min and then rinsed extensively with deionized water. Pre-dorsal closure stage embryos were selected using a reflected-light dissecting microscope. Embryos were prepared for imaging, as described previously (Kiehart et al., 1994, 2006). Images were acquired using Micro-Manager 2.0 software (Open Imaging) to control a Zeiss Axiovert 200 M microscope equipped with a Yokogawa CSU-W1 spinning disk confocal head (Solamere Technology Group), a Hamamatsu Orca Fusion BT camera and a Zeiss 40X LD LCI Plan-Apochromat 1.2NA multi-immersion objective (glycerin). Owing to the curvature of the embryo, we imaged multiple z planes for each embryo at each time point to view the dorsal opening. We recorded eight z-slice stacks with 1 μm step size for single color, and 14 z-slice stacks with 0.5 μm step size for dual-color movies. Stacks were acquired every 15 s throughout the duration of closure with a 100 ms exposure per slice for GFP and a 150 ms exposure per slice for mTomato.

Zebrafish husbandry and sample preparation for live imaging
Zebrafish of the Ekkwill strain were maintained between 26 and 28.5°C with a 14:10 h light:dark cycle. Fish between 3-6 months were used for experiments. Transgenic *krt4:lyn-EGFP* fish have been described previously (Lee et al., 2014). Male and female fish were set up for mating in tanks with dividers. The dividers were removed in the morning for timed mating. Embryos were collected in E3 medium and screened at 1 dpf for expression of GFP in the periderm (krt4-lynGFP). The positive embryos were transferred to a dish with E3 medium and tricaine (Sigma E10521-50G) at 0.01% concentration. The embryos were dechorionated with forceps and mounted in fluorinated ethylene propylene (FEP) tubes according to published protocols (Weber et al., 2014). The FEP tubes were coated with 3% methylcellulose and embryos were mounted in 0.1% agarose with 0.01% tricaine to immobilize them during imaging. The tube was then placed in a 60 mm culture dish with an agarose bed, held in place with 1% agarose and immersed in E3 medium. Images were acquired with LASX software on a Leica SP8 confocal microscope using an HC Fluotar L 25×/0.95NA W VISIR water-immersion objective at 0.75 or 1× zoom. Image stacks with 40-60 slices were acquired every 15 min with a z-step size of 2 μm. Work with zebrafish was approved by the Institutional Animal Care and Use Committee at Duke University.

Convolutional neural network architecture
DP uses an encoder-decoder convolutional neural network (Fig. 1B), inspired by the U-Net architecture (Falk et al., 2019). The left branch of the neural network extracts high-dimensional features by pairs of 3D convolutions with kernel size 3×3×3, each followed by Scaled Exponential Linear Units (SELU) activation. Between each double-convolutional layer, the content is down-sampled by max-pooling with kernel size 1×2×2 in only

DEVELOPMENT

the x and y directions. The number of convolution kernels further doubles in each layer. This multi-layer structure ensures the efficient extraction of both high-frequency features (such as bright cell boundary pixels) and low-frequency image features (such as whole cells in a tissue context). After feature extraction, the feature map is decoded and upscaled again to the initial stack dimensions using up-convolutions with kernel size $1{\times}2{\times}2$ alternating with 3D convolutions with kernel size $3{\times}3{\times}3$. To preserve the spatial resolution of shapes and boundaries, the input of each decoding layer is concatenated with the output of the corresponding encoding layer. After the last up-convolution layer, the output is scaled between 0 and 1 with sigmoid activation, yielding binary masks (Fig. 1C).

Training data generation and CNN model training

The ground truth (GT) used for training was created manually by human experts cutting out unwanted content (yolk granules, noise background, blurry out-of-focus cell boundaries and off-target tissue layers) of each image slice and keeping only the structure of interest (target epithelial tissue with sharp and distinctive tissue boundaries) with the freehand selection tool in Fiji/ImageJ (Schindelin et al., 2012) (Fig. 1F). This manual annotation procesure requires about 5-10 min per stack and only needs to be performed once. We selected not only the cell boundaries, but extended tissue regions containing whole cells. The masked stacks were binarized by clipping all remaining parts to 1. Our training dataset contained stacks of varying image quality and three different labeling strategies, with different image resolution, recorded with different microscopes (160 stacks for *Drosophila*, 20 stacks for zebrafish). As annotation of training data is a time consuming effort, we confirmed that a smaller training dataset of 5-10 is sufficient to obtain good result if the dataset is uniform, e.g. one fly line recorded on one microscope (Fig. S2; see supplementary Materials and Methods). We further tested whether this simple model can be directly transferred to distinct types of dorsal-closure data not contained in the small training dataset, from a different fly line recorded on a different microscope. We achieved only slightly inferior performance compared with both the generalist model (160 stacks of diverse conditions) and a simple model (five stacks) trained on the new data type. Furthermore, performance on new data types can be improved with reduced annotation effort by adapting an existing model towards new data types by training with a small additional training dataset.

The training dataset was augmented sixfold by randomly flipping and adjusting brightness and contrast. We aimed for fully binary masks with sharp and straight edges and contiguous regions. For training, we chose a log-cosh-Tversky loss function that yields sharp and distinctive mask edges (Nasalwai et al., 2021). This loss function can be tuned to punish false positives (high α, low β) or false negatives (low α, high β), and is identical to the common Dice loss for $\alpha{=}\beta{=}0.5$. We found optimal training results for $\alpha{=}0.3$ and $\beta{=}0.7$. We trained two separate networks, one for *Drosophila* and one for zebrafish, in each case for 50 epochs with learning rate 1e-5, stack patches of $(512{\times}512)$ pixels in x, y and batch size 12 on a workstation with a NVidia GTX 1080 Ti GPU.

Uniaxial flattening of curved tissue sheets

To uniaxially flatten curved tissue sheets, unique z-height maps are created by averaging the positive z-indices of the binary masks predicted by DP. Then the z-height maps are smoothed by 2D averaging with a kernel size of $3{\times}3$ pixels (Fig. 4H). The gradient tensor $\alpha\,(i,j)$ of the z-height maps then yields the local curvature in x and y. This makes it possible to locally correct the distortion for individual cells using an affine transformation with a gradient tensor of the z-height map averaged over the proximity of each cell (not shown here). Alternatively, the tissue can be flattened in only one direction. This is particularly useful for tube-like tissues. First, a marker line at position k is defined manually. Then, the curved tissue is cut in one-pixel wide stripes, straightened, stitched back together and the stripes aligned at the previously defined marker line k. The unidirectional transformation map $T\,(i,j)$ for each pixel (i,j) can thus be defined as:

$$T_x(i,j) = \sum_{n=k}^{i} \frac{i-k}{|i-k|}\left(\sqrt{1 + \alpha_x(n,j)^2}\right) - (i-k).$$

DeepProjection code, hardware requirements and data availability

DeepProjection is implemented in Python 3.8 using standard free packages: numpy 1.19 for scientific computing, pytorch 1.7.1 (with cuda 11.0) for deep learning, albumentations 0.5.2 for data augmentation (other versions might also work, but not guaranteed). The DeepProjection repository (https://github.com/danihae/DeepProjection/) contains the DP code and Jupyter notebooks with detailed instructions for training, prediction and flattening. Additionally, we created a graphical user interface for prediction and batch processing (Fig. S3). Training data, trained DP models and the test data used for benchmarking for *Drosophila* dorsal closure and *Danio* periderm development are available from the Dryad Digital Repository (Haertter et al., 2022): dryad.x0k6djhnf. The DeepProjection package is further available on Python Package Index (PyPI) (https://pypi.org/project/deepprojection/).

DeepProjection was trained on a workstation with 32 GB onboard memory and NVIDIA GeForce 1080 Ti with Windows 10 operating system. Training and prediction were also successfully tested on regular desktop computers and laptops without dedicated GPU with Windows 10 and Ubuntu 20.04 operating system.

Competing interests
The authors declare no competing or financial interests.

Author contributions
Conceptualization: D.H., D.P.K., C.F.S.; Methodology: D.H., X.W.; Software: D.H., X.W.; Validation: D.H., X.W.; Formal analysis: D.H., X.W.; Investigation: D.H., X.W., N.R., J.M.C., S.M.F.; Resources: K.D.P., S.D.T., D.P.K., C.F.S.; Data curation: D.H., X.W., N.R., J.M.C., S.M.F.; Writing - original draft: D.H.; Writing - review & editing: D.P.K., C.F.S.; Visualization: D.H.; Supervision: D.P.K., C.F.S.

Data availability
Data are available from the Dryad Digital Repository (Haertter et al., 2022): dryad.x0k6djhnf.

Peer review history
The peer review history is available online at https://journals.biologists.com/dev/lookup/doi/10.1242/dev.200621.reviewer-comments.pdf.

References
Basu, S., Rexhepaj, E., Spassky, N., Genovesio, A., Paulsen, R. R. and Shihavuddin, A. S. M. (2018). FastSME: faster and smoother manifold extraction from 3D stack. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit* 2281-2289. doi:10.1109/CVPRW.2018.00305

Belthangady, C. and Royer, L. A. (2019). Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction. *Nat. Methods* **16**, 1215-1225. doi:10.1038/s41592-019-0458-z

Berg, S., Kutra, D., Kroeger, T., Straehle, C. N., Kausler, B. X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M. et al. (2019). ilastik: interactive machine learning for (bio)image analysis. *Nat. Methods* **16**, 1226-1232. doi:10.1038/s41592-019-0582-9

Blasse, C., Saalfeld, S., Etournay, R., Sagner, A., Eaton, S. and Myers, E. W. (2017). PreMosa: extracting 2D surfaces from 3D microscopy mosaics. *Bioinformatics* **33**, 2563-2569. doi:10.1093/bioinformatics/btx195

Chang, W.-J. and Hwang, P.-P. (2011). Development of zebrafish epidermis. *Birth Defects Res. C Embryo Today* **93**, 205-214. doi:10.1002/bdrc.20215

Duque, J. and Gorfinkiel, N. (2016). Integration of actomyosin contractility with cell-cell adhesion during dorsal closure. *Development* **143**, 4676-4686. doi:10.1242/dev.136127

DEVELOPMENT

Eisenhoffer, G. T., Slattum, G., Ruiz, O. E., Otsuna, H., Bryan, C. D., Lopez, J., Wagner, D. S., Bonkowsky, J. L., Chien, C.-B., Dorsky, R. I. et al. (2017). A toolbox to study epidermal cell types in zebrafish. *J. Cell Sci.* **130**, 269-277. doi:10.1242/jcs.184341

Erguvan, Ö., Louveaux, M., Hamant, O. and Verger, S. (2019). ImageJ SurfCut: a user-friendly pipeline for high-throughput extraction of cell contours from 3D image stacks. *BMC Biol.* **17**, 38. doi:10.1186/s12915-019-0657-1

Falk, T., Mai, D., Bensch, R., Çiçek, Ö., Abdulkadir, A., Marrakchi, Y., Böhm, A., Deubner, J., Jäckel, Z., Seiwald, K. et al. (2019). U-Net: deep learning for cell counting, detection, and morphometry. *Nat. Methods* **16**, 67-70. doi:10.1038/s41592-018-0261-2

Forster, B., Ville, D. V. D., Berent, J., Sage, D. and Unser, M. (2004). Complex wavelets for extended depth–of–field: A new method for the fusion of multichannel microscopy images. *Microsc. Res. Tech* **65**, 33-42. doi:10.1002/jemt.20092

Haertter, D., Wang, X., Fogerson, S. M., Ramkumar, N., Crawford, J. M., Poss, K. D., Di Talia, S., Kiehart, D. P. and Schmidt, C. F. (2022). Data from: DeepProjection: Specific and robust projection of curved 2D tissue sheets from 3D microscopy using deep learning. *Dryad Digital Repository* doi:10.5061/dryad.x0k6djhnf

Heemskerk, I. and Streichan, S. J. (2015). Tissue cartography: compressing bio-image data by dimensional reduction. *Nat. Methods* **12**, 1139-1142. doi:10.1038/nmeth.3648

Herbert, S., Valon, L., Mancini, L., Dray, N., Caldarelli, P., Gros, J., Esposito, E., Shorte, S. L., Bally-Cuif, L., Aulner, N. et al. (2021). LocalZProjector and DeProj: a toolbox for local 2D projection and accurate morphometrics of large 3D microscopy images. *BMC Biol.* **19**, 136. doi:10.1186/s12915-021-01037-w

Huang, J., Zhou, W., Dong, W., Watson, A. M. and Hong, Y. (2009). Directed, efficient, and versatile modifications of the Drosophila genome by genomic engineering. *Proc. Natl. Acad. Sci. USA* **106**, 8284-8289. doi:10.1073/pnas.0900641106

Kiehart, D. P., Crawford, J. M., Aristotelous, A., Venakides, S. and Edwards, G. S. (2017). Cell sheet morphogenesis: dorsal closure in drosophila melanogaster as a model system. *Annu. Rev. Cell Dev. Biol.* **33**, 169-202. doi:10.1146/annurev-cellbio-111315-125357

Kiehart, D., Tokutake, Y., Chang, M.-S., Hutson, M., Weimann, J., Peralta, X., Toyama, Y., Wells, A., Rodriguez, A. and Edwards, G. (2006). Ultraviolet laser microbeam for dissection of Drosophila embryos. In *Cell Biology: A Laboratory Handbook*, Vol. 3 (ed. J. E. Celis), Chapter 9. New York: Elsevier Academic Press.

Kiehart, D. P., Galbraith, C. G., Edwards, K. A., Rickoll, W. L. and Montague, R. A. (2000). Multiple forces contribute to cell sheet morphogenesis for dorsal closure in Drosophila. *J. Cell Biol* **149**, 471-490. doi:10.1083/jcb.149.2.471

Kiehart, D. P., Montague, R. A., Rickoll, W. L., Foard, D. and Thomas, G. H. (1994). Chapter 26 high-resolution microscopic methods for the analysis of cellular movements in Drosophila Embryos. *Methods Cell Biol.* **44**, 507-532. doi:10.1016/S0091-679X(08)60929-2

Lee, R. T. H., Asharani, P. V. and Carney, T. J. (2014). Basal keratinocytes contribute to all strata of the adult Zebrafish epidermis. *PLoS ONE* **9**, e84858. doi:10.1371/journal.pone.0084858

Ma, X., Lynch, H. E., Scully, P. C. and Hutson, M. S. (2009). Probing embryonic tissue mechanics with laser hole drilling. *PhBio* **6**, 036004. doi:10.1088/1478-3975/6/3/036004

Mavrakis, M., Rikhy, R., Lilly, M. and Lippincott-Schwartz, J. (2008). Fluorescence imaging techniques for studying drosophila embryo development. *Curr. Protoc. Cell Biol.* **39**, 4.18.1-4.18.43. doi:10.1002/0471143030.cb0418s39

Nasalwai, N., Punn, N. S., Sonbhadra, S. K. and Agarwal, S. (2021). Addressing the class imbalance problem in medical image segmentation via accelerated tversky loss function. *Lect. Notes Comput. Sci* **12714**, 390-402. doi:10.1007/978-3-030-75768-7_31

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid B. et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676-682. doi:10.1038/nmeth.2019

Sokolow, A., Toyama, Y., Kiehart, D. P. and Edwards, G. S. (2012). Cell ingression and apical shape oscillations during dorsal closure in Drosophila. *Biophys. J* **102**, 969-979. doi:10.1016/j.bpj.2012.01.027

Weber, M., Mickoleit, M. and Huisken, J. (2014). Multilayer mounting for long-term light sheet microscopy of zebrafish. *J. Vis. Exp* **84**, e51119. doi:10.3791/51119

Weigert, M., Schmidt, U., Boothe, T., Müller, A., Dibrov, A., Jain, A., Wilhelm, B., Schmidt, D., Broaddus, C., Culley, S. et al. (2018). Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat. Methods* **15**, 1090-1097. doi:10.1038/s41592-018-0216-7

DEVELOPMENT