

MMV_IM2IM: An Open Source Microscopy Machine Vision Toolbox for Image-to-Image Transformation

This manuscript ([permalink](#)) was automatically generated from [MMV-Lab/im2im-paper@d59eb7e](https://gitlab.com/MMV-Lab/im2im-paper@d59eb7e) on September 4, 2022.

Authors

- **Justin Sonneck**

 [0000-0002-1640-3045](#) ·  [Justin-Sonneck](#) ·  [JustinSonneck](#)

Leibniz-Institut für Analytische Wissenschaften - ISAS - e.V. · Funded by The Federal Ministry of Education and Research (BMBF) under the funding reference 161L0272

- **Jianxu Chen**

 [0000-0002-8500-1357](#) ·  [jxchen01](#) ·  [JianxuChen](#)

Leibniz-Institut für Analytische Wissenschaften - ISAS - e.V. · Funded by The Federal Ministry of Education and Research (BMBF) under the funding reference 161L0272

Abstract

Introduction

With the fast development of the machine learning (ML) and computer vision fields, computers are now able to transform images into new forms for better visualization [1], better animation [2], better information extraction [3] with unprecedented and continuously growing accuracy and efficiency. Recently, such techniques have started to be adapted for bioimaging applications and revolutionized image-based biomedical research [4,5,6,7]. In principle, these techniques and applications can be formulated as a general image-to-image transformation problem, as illustrated in the central panel in Figure 1. In essence, deep neural networks are trained to perceive the information from the source image(s) and reconstruct the learned knowledge from source images(s) in the form of a new image(s) of the target type. Here, the source and target images can be real microscopy images, simulated microscopy images, segmentation masks, or their combinations, such as the examples in Figure 1. Considering the common essential spirit of all the underlying methods, a natural question to ask is whether it is possible to build a single generic codebase for deep learning based image-to-image transformation directly applicable to different biomedical studies?

In this paper, we introduce *MMV_Im2Im* an open-source microscopy machine vision (MMV) toolbox for image-to-image transformation that can be used in various biomedical applications. Currently, *MMV_Im2Im* supports 2D~5D microscopy images for supervised image to image translation (e.g., labelfree determination [4], imaging modality transformation [5,8]), supervised image restoration [6], supervised semantic segmentation [9], supervised instance segmentation [10], unsupervised semantic segmentation [11], unsupervised image to image translation and synthetization [12]. The toolbox will continuously grow and more methods will be supported, especially methods based on self-supervised learning.

Why do we need such a generic codebase for all deep-learning based microscopy image-to-image transformation? *MMV_Im2Im* is not simply a collection of many existing methods, but with rather systematic design for generality, flexibility and simplicity, attempting to address many fundamental pain points for image-to-image transformation in biomedical applications, as highlighted below.

Universal boilerplate with state-of-the-art ML engineering:

The toolbox employs pytorch-lightning [13] as the core in the backend, which offers numerous benefits, such as readability, flexibility and simplicity. First of all, have you ever encountered the situation where you want to understand the code from two different papers, even solving the same problem, e.g. semantic segmentation, but not quite easy to grasp quickly since the two repositories are implemented in very different ways? It is not rare that even different researchers from the same group may implement similar methods in very different manners. This is not only a barrier for other people to learn and re-use the open-source code, but also poses challenges for developers in maintenance, further development, and interoperability among different packages. We follow the pytorch-lightning framework and carefully design a universal boilerplate for image-to-image transformation for biomedical applications, where the implementation of all the methods share the same modularized code structure. This greatly lowers the learning curve for people to read and understand the code, and makes implementing new methods or extending existing methods simple and fast.

Moreover, as ML scientists, have you ever overwhelmed by different training tricks for different methods or been curious about if certain state-of-the-art training methods can boost the performance of your models? With the pytorch-lightning backend, *MMV_Im2Im* allows you to enjoy

different state-of-the-art ML engineering techniques without changing any line of code, e.g., stochastic weight averaging [14], single precision training, automatic batch size determination, different optimizers, different learning rate schedulers, easy deployment on different devices, distributed training on multi-GPU (even multi-node), logging with common loggers such as Tensorboard, etc. In short, with the pytorch-lightning based universal boilerplate, one can really focus on research and develop novel methods for bioimaging applications, without worrying on the ML engineering works (which are usually lack in non-computer-science labs).

Modularization and human-readable configuration system:

The toolbox is designed for both computational biomedical imaging researchers (e.g., with expertise in biomedical imaging but only basic knowledge of Python or ML) and ML researchers (e.g. deep knowledge of ML methodology but with limited experience in microscopy). For this purpose, we design the toolbox in a systematically modularized way with various levels of configurability. One can use the toolbox with a single command as simple as `run_im2im --config train_semanticseg_3d --data.data_path /path/to/data` or make customization on details directly from a human-readable configuration file, such as choosing batch normalization or instance normalization in certain layers of the model, or adding extra data augmentation steps, etc. For users without experience in Python programming, another MMV toolbox has been planned as the extension of *MMV_Im2Im* (See the Discussion section for details). In addition, the modularization and configuration system is designed to allow not only configuring with the elements offered by the package itself, but also any compatible elements from a third-party package or from a public repository on Github. For example, one can easily switch the 3D neural network in the original *Embedseg* method to any customized U-Net from FastAI by specifying the network as `fastai.vision.models.unet`. Such painless extendability releases the power of the toolbox, amplifies the benefit of the open-source ML community and upholds our philosophy of open science.

Customization for biomedical imaging applications:

The original idea of a general toolbox actually stems from the OpenMMLab project, which provides generic codebase for a wide range of computer vision research topics. For instance, MM Segmentation is an open source toolbox for semantic segmentation, supporting unified benchmarking and state-of-the-art models ready to use out-of-box. It has become one of most widely used codebase for research in semantic segmentation (1.6K forks and 4.3K stars on Github as of August 30, 2022). This inspires us to develop *MMV_Im2Im* to fascinate research in general image-to-image transformation with special focus on biomedical applications.

First of all, different from general computer vision datasets, such as ImageNet [15], where the images are usually small 2D RGB images (e.g., 3 x 256 x 256 pixels), bioimaging applications usually involves large-scale high dimensional data (e.g., 4 x 128 x 2048 x 2048 voxels). To deal with this issue, we employ the PersistentDataset in MONAI [16] with partial loading and sampling support, as well as delayed image reading in aicsimageio [17]. As a result, in our stress test, training an 3D nuclei instance segmentation model with more than 125,000 3D images can be conducted efficiently, even with limited resource.

Second, because microscopy data is not restricted to 2D, we re-implement common frameworks, such as fully convolutional networks (FCN), conditional generative models, cycle-consistent generative models, etc., in a generic way to easily switch between different dimensionalities.

Third, the toolbox pre-packs common functionalities specific to microscopy images. For example, we incorporate the special image normalization method introduced in [4], where only the middle chunk along Z dimension of 3D microscopy images will be used for calculating the mean and standard

deviation of image intensity for standard normalization. Also, 3D light microscopy images are usually anisotropic, i.e., much lower resolution along Z than XY dimension. So, we adopt the anisotropic variation of UNet as proposed in [18].

Finally, to deploy the model in production, a model trained on small 3D patches sometimes need to be applied not only on much large images, but also with additional dimensionalities (e.g., multi-scene timelapse). Combining the efficient data handling of aicsimageio [17] and the sliding window inference with gaussian weighted blending, the toolbox can yield efficient inference without visible stitching artifacts in production.

All in all, the *MMV_Im2Im* toolbox stands on the shoulders of many giants in the open-source software and ML engineering communities (pytorch-lightning, MONAI, aicsimageio, etc.) and is systematically designed for image-to-image transformation R&D for biomedical applications. The source code of **MMV_Im2Im** is available at https://github.com/MMV-Lab/mmv_im2im. This manuscript is generated with open-source package Manubot [19]. The manuscript source code is available at <https://github.com/MMV-Lab/manuscript-im2im>.

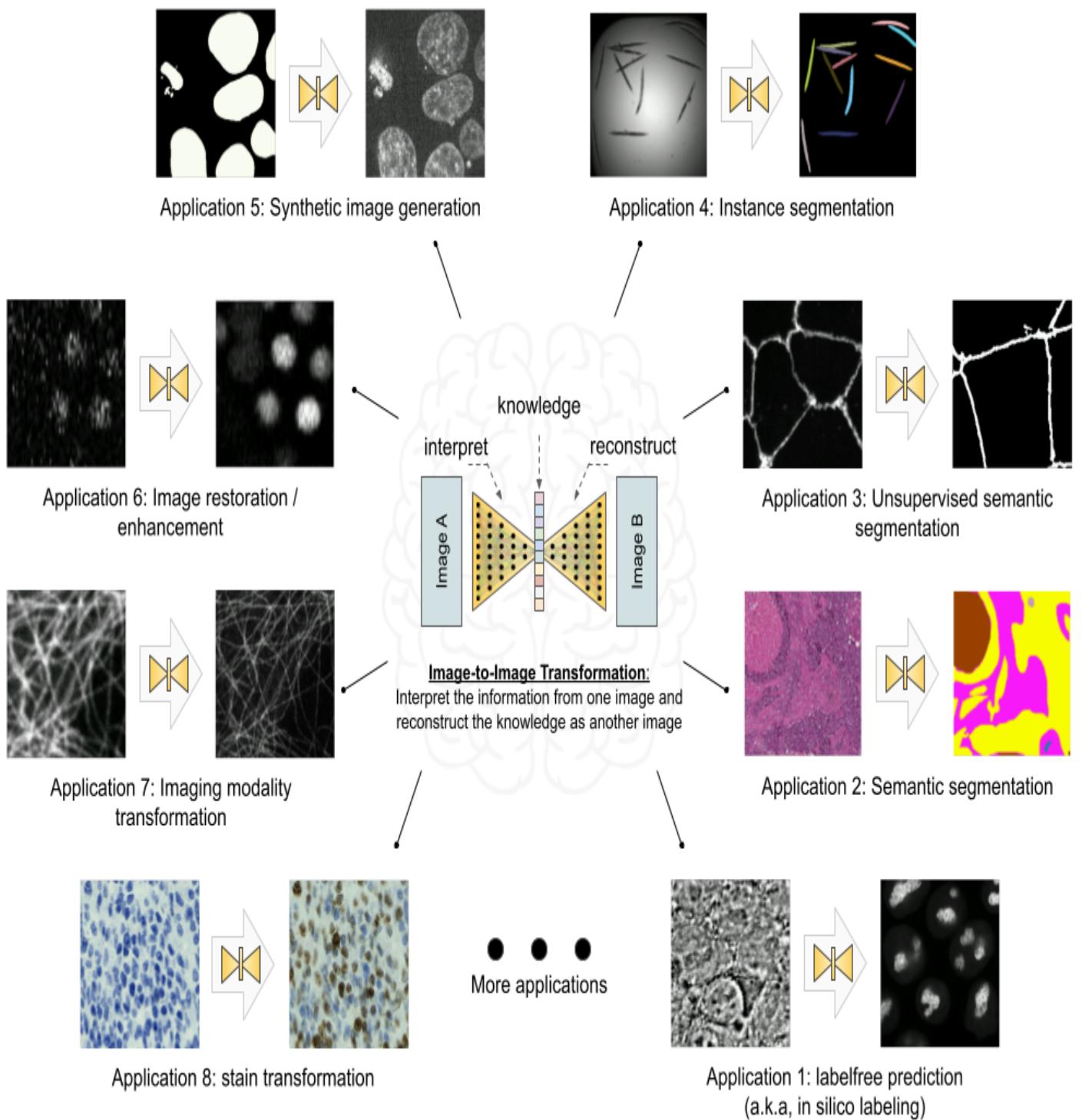


Figure 1: Overview of the image-to-image transformation concept and its example applications

Results

In this section, we demonstrated the versatility of the **MMV_Im2Im** toolbox with over 10 different biomedical applications from different R&D use cases at different scales. All experiments and results in this section were conducted on public datasets and can be reproduced with the scripts (for pulling the public dataset online and data wrangling) and configuration files (for setting up training and inference details) released with the **MMV_Im2Im** package. It is worth emphasizing that the aim of these experiments were not to achieve the state-of-the-art performance on each individual task (may require further tuning of hyper-parameters, see Discussion section for more details). All experiments were conducted without much tuning, and essentially meant to demonstrate different features of the package and the general applicability. We intend to provide a holistic view of the image-to-image transformation concepts to biomedical researchers using a diverse set of applications. Hopefully, such

concepts could help biomedical researchers to better integrate AI into traditional assay development strategies and inspire more computational and experimental co-design methods to realize new biomedical studies that were not very feasible before.

Labelfree prediction of nuclear structure from 2D/3D brightfield images

The labelfree method refers a deep learning method that can predict fluorescent images directly from transmitted light brightfield images [4]. Comparing to brightfield images, fluorescent images can resolve subcellular structures in living cells at high resolution but with the cost of expensive and slow procedures and high phototoxicity. The labelfree method provides a new perspective in assay development to conduct integrated computational analysis of multiple organelles only with a single brightfield imaging acquisition. In our first demonstration, we applied **MMV_Im2Im** to build 2D/3D models that can predict fluorescent images of nuclear structures from brightfield images. For 3D models, we also compared (1) different image normalization methods, (2) different network backbones, and (3) different types of models.

It is worth mentioning that systematic evaluation of the prediction is beyond the scope of this paper. We argue that a proper evaluation depends on specific downstream quantitative analysis goals. For example, if we want to quantify the size of the nucleoli, we have to compare the segmentation from real nucleoli signals and the segmentation from predicted nucleoli, and make sure the sizes measured from real signal are consistent with measurement from predictions. In contrast, if the goal is to roughly localize the nucleoli within the cell, pearson correlation might be more suitable. As a result, we only compare the results by visual inspection, for the sake of simplicity and focusing on the applicability of our **MMV_Im2Im** package, and leave the systematic evaluations to users in specific problems in real studies.

2D Model: We started with a simple problem using 2D images from the HeLa "Kyoto" cells dataset [20]. For all images, we took the brightfield channel and the mCherry-H2B channel out of the multi-channel timelapse movies. 2D images were acquired at 20x with 0.8 N.A. and then downsampled by 4 (pixel size: 0.299 nm x 0.299 nm). Example predictions can be found in Figure 2-A. We compared a basic UNet model [9] and a 2D version of the fnet model in [4]. The fnet model achieved slightly more accurate predictions.

3D Model: We tested with 3D images from the hiPS single cell image dataset [21]. Specifically, we extracted the brightfield channel and the structure channel from the full field-of-view (FOV) multi-channel images, from the HIST1H2BJ, FBL, NPM1, LMNB1 cell lines, so as to predict from one brightfield image various nuclear structures, histones, nucleoli (dense fibrillar component via fibrillarin), nucleoli (granular component via nucleophosmin), and nuclear envelope, respectively. Images were acquired at 100x with 1.25 NA. (voxel size: 0.108 micron x 0.108 micron x 0.29 micron).

We conducted three groups of comparisons (see results in Figure 2). First, we compared three different image normalization methods for 3D images (percentile normalization, standard normalization, center normalization [4]). Percentile normalization refers to cutting the intensity out of the range of [0.5, 99.5] percentile of the image intensity and then rescale to the range of [-1, 1], while the standard normalization is simply subtracting mean intensity and then divided by the standard deviation of all pixel intensities. Center normalization is similar to standard normalization, but the statistics are calculated only near around center along the Z-axis. One could easily test different percentile or rescaling to [0, 1] instead of [-1, 1]. Qualitatively, we found center normalization slightly more accurate and more robust than the other two (ref. the first row in Figure 2-B).

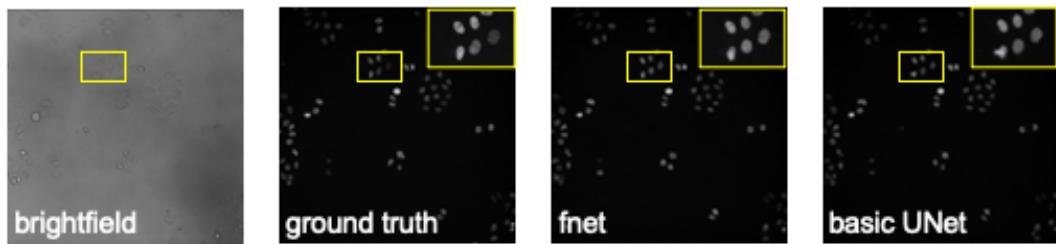
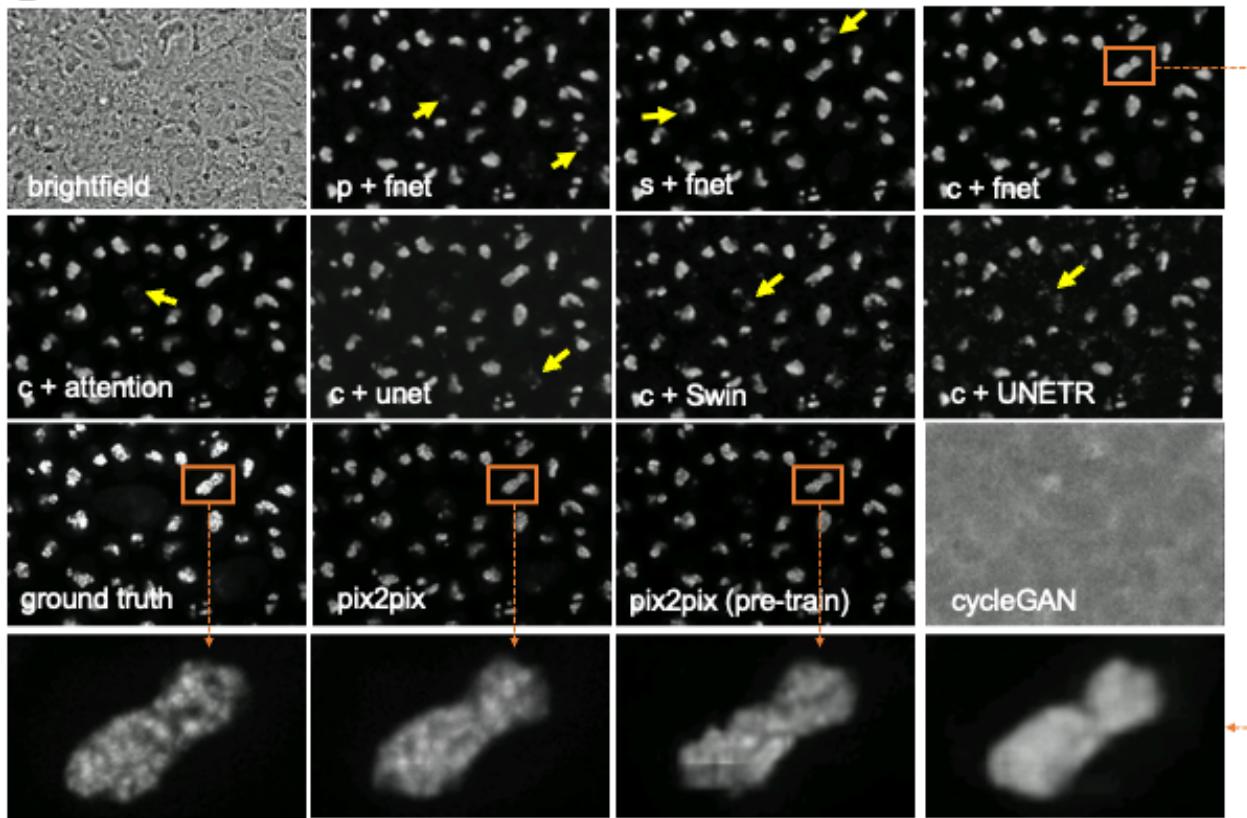
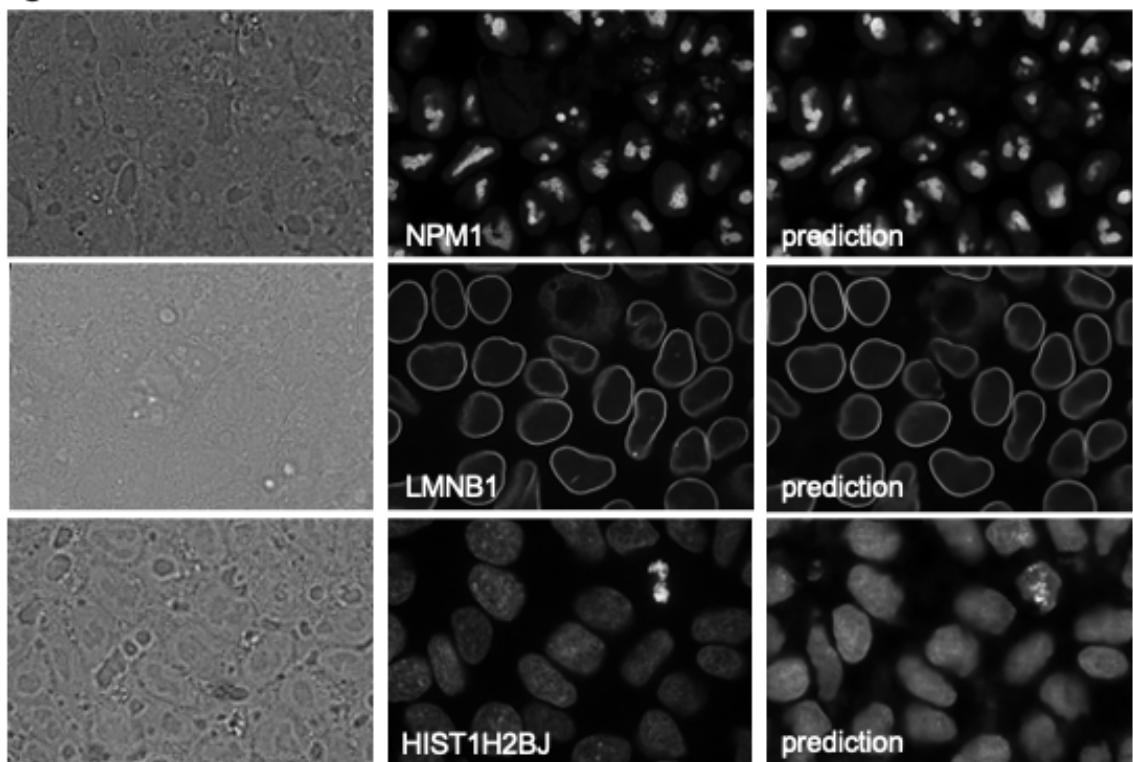
Second, we compared different network backbone architectures, including original fnet model [4], enhanced UNet [22], attention UNet [23], two transformer-based models, SwinUNETR [24] and UNETR[25] (all with center normalization). Inspecting the predictions on a holdout validation set suggested that fnet achieved the best performance, and the selected transformer-based models did not work well with in labelfree problems (ref. the second row and the “c + fnet” from the first row in Figure 2-B).

Finally, we showed the comparison between three different types of models, an FCN-type model (i.e., fnet), a pix2pix-type model, and a cycleGAN-type model. For fair comparison, we used fnet as the generator in the pix2pix-type model, while coupling with the multi-scale discriminator from pix2pixHD [26]. Then, the pix2pix-type model can be trained in two different ways: from scratch or initializing the generator with a pre-trained fnet. Examples of the comparison results are shown in the last two rows in Figure 2-B. Visually, it is evident that the additional adversarial components (i.e., the discriminator) could boost the performance of a typical FCN-type model to generate more realistic images than an FCN-type model alone.

From the experiments above, we found that center normalization + pix2pix with fnet as the generator achieved the best performance. So, we employed the same strategy on all other nuclear structures. At the end, we had four different labelfree models, each predicting one different nuclear structure from 3D brightfield images. As an example of evaluation, we calculated the pearson correlation and structural similarity on the validation set. The results were summarized in Table 1. Again, these numbers are merely examples of evaluation, systematic evaluation based each specific biological problem is necessary before deployment. Figure 2-C shows one example of all four different structures predicted from a single unseen brightfield image. This would permit an integrated analysis of cell nuclear components that could hardly be done with real experiments and real images.

Table 1: Evaluation of the final 3D label-free models for the four different nuclear structures.

Dataset	Pearson Correlation	Structural Similarity	# of Test Data
FBL	0.864 &pm 0.021	0.831 &pm 0.034	50
HIST1H2BJ	0.825 &pm 0.034	0.675 &pm 0.073	55
LMNB1	0.853 &pm 0.027	0.669 &pm 0.059	50
NPM1	0.912 &pm 0.015	0.795 &pm 0.039	55

A**B****C****D**

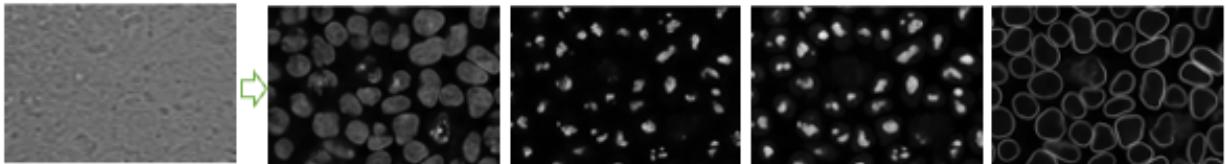
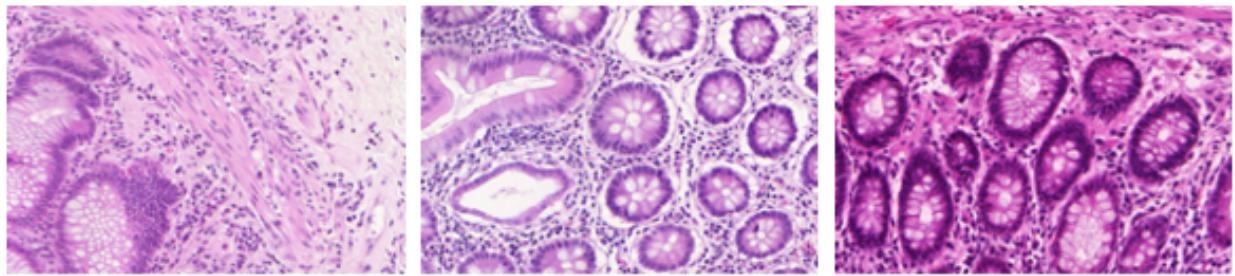


Figure 2: Examples of labelfree results.

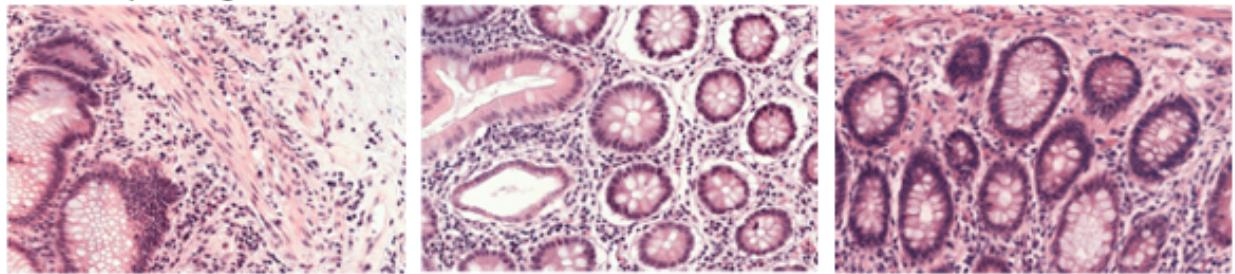
2D semantic segmentation of tissues from H&E images

Segmentation is a common image processing task, and can be considered as a special type of image-to-image transformation, where the generated images are segmentation masks. Deep learning based methods have achieved huge success in 2D semantic segmentation in biomedical images. In this example, we demonstrate **MMV_Im2Im** on a pathology application to segment glands from hematoxylin and eosin (H&E) stained tissue images from the 2015 Gland Segmentation challenge [27,28]. Stain normalization is an important pre-processing step in order to develop models robust to stain variation and tissue variations. **MMV_Im2Im** included a classic stain normalization method [29] as a pre-processing step. The effect of stain normalization can be observed in Figure 3-A and B. We trained a simple attention UNet model [23]. Evaluated on the two different hold-out test sets, the model achieved F1-score, 0.883 and 0.888 on test set A and test set B, respectively. The performance is competitive comparing to the methods reported in [28], especially with much more consistent performance across the two different test sets. Example results can be found in Figure 3-C.

A: Example images (original)



B: Example images after stain normalization



C: Example results on two different test sets

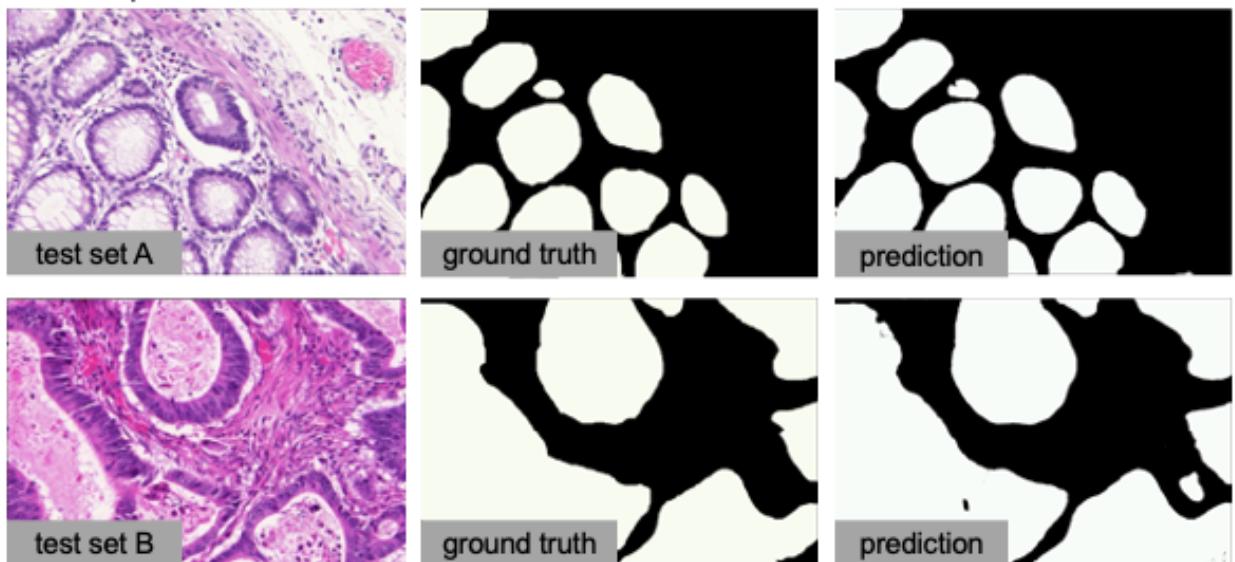


Figure 3: Example results of 2D semantic segmentation of gland in H&E images.

3D semantic segmentation of organelles from electron microscopy images

Semantic segmentation in 3D biomedical image analysis application is not a simple generalization from 2D models by switching 2D operations with 3D operations, but with many practical challenges. Large GPU footprint is one of the biggest issues, which makes many training strategies common in 2D not feasible in 3D, e.g. limited mini-batch size. **MMV_Im2Im** is able to take advantage of state-of-the-art ML engineering methods to efficiently handle 3D problems. For example, by using effective half-precision training, we are able to greatly reduce GPU memory workload for each sample and therefore increase the batch size. When multiple GPUs are available, we can easily take advantage of the additional resources to scale up the training to multiple GPU cards, even multiple GPU nodes. Here, we trained a 3D model to segment seven different types of organelles (e.g., cell, mitochondrion, alpha granule, etc.) from SBF-SEM image volumes [30]. Example results can be found in Figure 4.

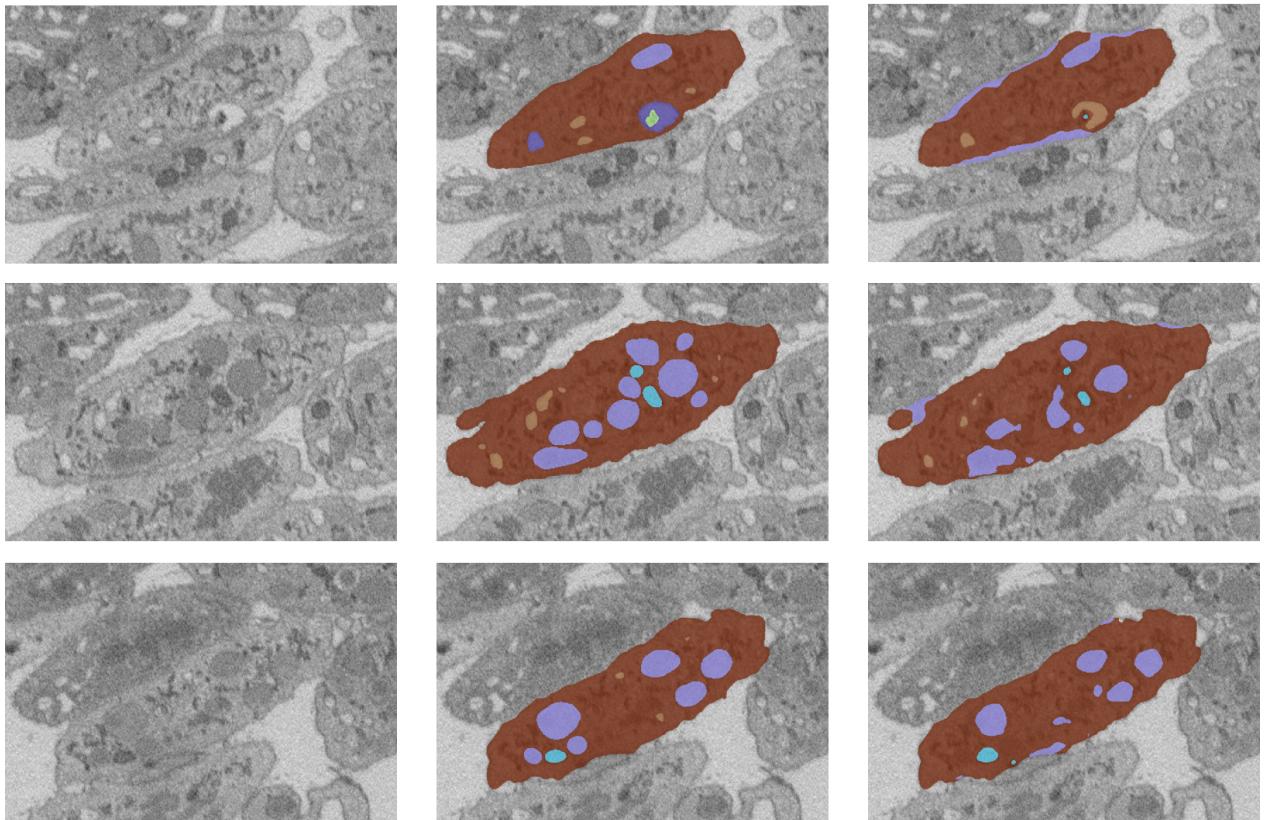


Figure 4: Result of 3D semantic segmentation results on z-slice 30, 60, 90. From left to right: raw images, ground truth, predictions

Unsupervised semantic segmentation of intracellular structures from 2D/3D confocal microscopy images

Large amount of high-quality training ground truth is not always available or may take endless effort to collect. CycleGAN based methods opened a new venue for segmentation without pixel-wise ground truth [11]. We demonstrated unsupervised learning based segmentation method for four examples: 2D tight-junction (via ZO1) segmentation from 2D FP-tagged ZO1 images (max-projected from 3D stacks) and segmentation of nuclei, mitochondria and golgi from 3D confocal microscopy images. Here, we took raw images from the hiPS single cell image dataset [21] and their corresponding segmentation. We shuffled the raw images and their segmentation to perform unsupervised learning. In theory, the segmentation masks could be simulated as suggested in [11]. Here, for the sake of simplicity in the demonstration, we took the original segmentation (shuffled) as a set of simulated segmentation masks. The concept is illustrated in Figure 5-A and example results of all 3D models are shown in Figure 5-B. The F1-scores of on test set were summarized in Table 2. According to [11], the quality of unsupervised nuclei segmentation could be further improved with additional simulation strategies. In general, we can observe that unsupervised learning could offer an effective way to generate preliminary segmentation, which can be further curated to train additional models in supervised ways, like the iterative deep learning workflow in [18].

Table 2: F1 scores of the unsupervised semantic segmentation predictions.

Dimensionality	Dataset	F1 Score	# of Test Data
2D	TJP1	0.888 &pm 0.022	29
3D	AAVS1	0.465 &pm 0.046	3
3D	HIST1H2BJ	0.811 &pm 0.150	15
3D	ST6GAL1	0.705 &pm 0.022	6

Dimensionality	Dataset	F1 Score	# of Test Data
3D	TOMM20	0.783 &pm 0.005	2

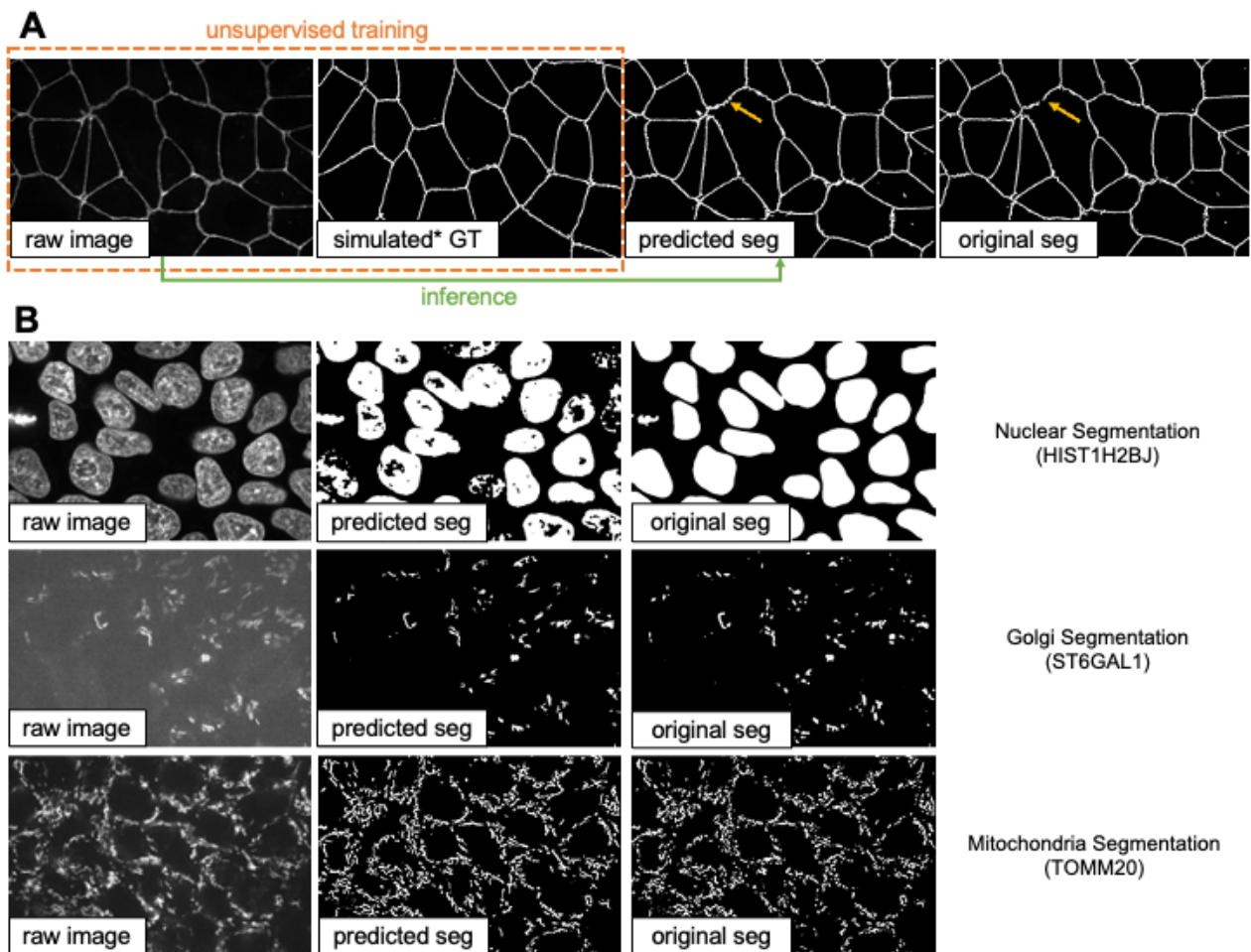


Figure 5: (A) Illustration of the unsupervised training scheme and results in the 2D tight-junction segmentation problem. (B) Example 3D segmentation results (only showing a middle z-slice) from models obtained by unsupervised training.

Instance segmentation in microscopy images

Another type of segmentation, other than semantic segmentation, is the instance segmentation problem. The goal is to not only differentiate different types of objects, but also different instances of the same type of objects. Currently, **MMV_Im2Im** supports *EmbedSeg*-type models (re-implemented from the original paper following the generic boilerplate in **MMV_Im2Im**). The major benefit of *EmbedSeg*-type models is that the method is agnostic to the morphology and dimensionality of the objects, comparing to other FCN-type models like StarDist [31,32] (difficult in handling different shapes, e.g. elongated objects) and CellPose [33] or SplineDist [34] (not straightforward to generalize to higher dimensions). Another category of mainstream instance segmentation methods is detection based models, such as Mask-RCNN [35], which fits more to the detection framework, instead of image-to-image transformation (See Discussion section for details). As a demonstration, we applied *EmbedSeg*-like models on a 2D problem of segmenting *C. elegans* from widefield images [36] and a 3D problem of nuclear segmentation from fluorescent images and brightfield images from the hiPS single cell image dataset [21].

For the 2D problem, we adopted the same network backbone as in the original *EmbedSeg* paper. Example results on a small holdout set are shown in Figure 6-A (IoU = 0.86). For the 3D problem, the original method is not directly applicable for two reasons. First, we cannot use the original network

architecture, ERF-Net-3D, because the images have very anisotropic dimensions and there are not enough Z-slices to run through all down sampler blocks in 3D. Second, we cannot use the original training function, because some parts of the images do not have ground truth and the original training method requires fully annotated images. (Note: the segmentation results from the public dataset [21] contain nuclear instance segmentation of all cells. But, the cells touch the image borders are ignored from downstream analysis and therefore not curated. In other words, the segmentation from this public dataset can be used as high-quality nuclear instance segmentation ground truth after excluding the areas covered by cells touching the image borders. See more details in [21].) In this example, we showcase (1) how to easily swap the network backbone with a network from public repositories on Github, and (2) how to easily train the model with pixel-wise cost map or exclusion masks (i.e., certain pixels can be excluded from training). Another extension comparing to the original implementation, the **MMV_Im2Im** package makes sliding windowing inference straightforward, and therefore permits easy handling of images of any size during inference in practice. For network, we directly imported a model from a public github repository <https://github.com/AllenCell/aics-ml-segmentation> for segmentation in fluorescent images and a model from another public Github repository https://github.com/AllenCellModeling/pytorch_fnet for segmentation in brightfield images. For exclusion masks, all the areas covered by cells touching the image border were assigned to 0, otherwise 1, as a cost map.

Example results are presented in 6-B. The green box highlight a mitotic cell, where the DNA signals form “spaghetti” shapes. Besides roughly separating the DNA signals from background, the model is also able to correctly identify the instance identity, which is theoretically infeasible for other FCN-type instance segmentation models. Nuclear instance segmentation from brightfield images is much more challenging than from fluorescent images. Arguably, this could be thought of as one single model doing two transformations: predicting the DNA signals from brightfield and running instance segmentation on predicted DNA signals. From Figure 6-B, it is shown that the segmentation from brightfield images is comparable to the segmentation from fluorescent images, but with two caveats. First, the performance on mitotic cells is worse than the model using fluorescent images. We hypothesize this is due to the limited information in brightfield images for mitotic cells, compounded with limited number of mitotic cells in the whole training set (much less than 10%). Second, the performance along Z dimension is also worse than the fluorescent model, as explained by the side view in Figure 6-B. This could be explained by the different properties of brightfield imaging and fluorescent imaging, but would need further comprehensive studies to investigate.

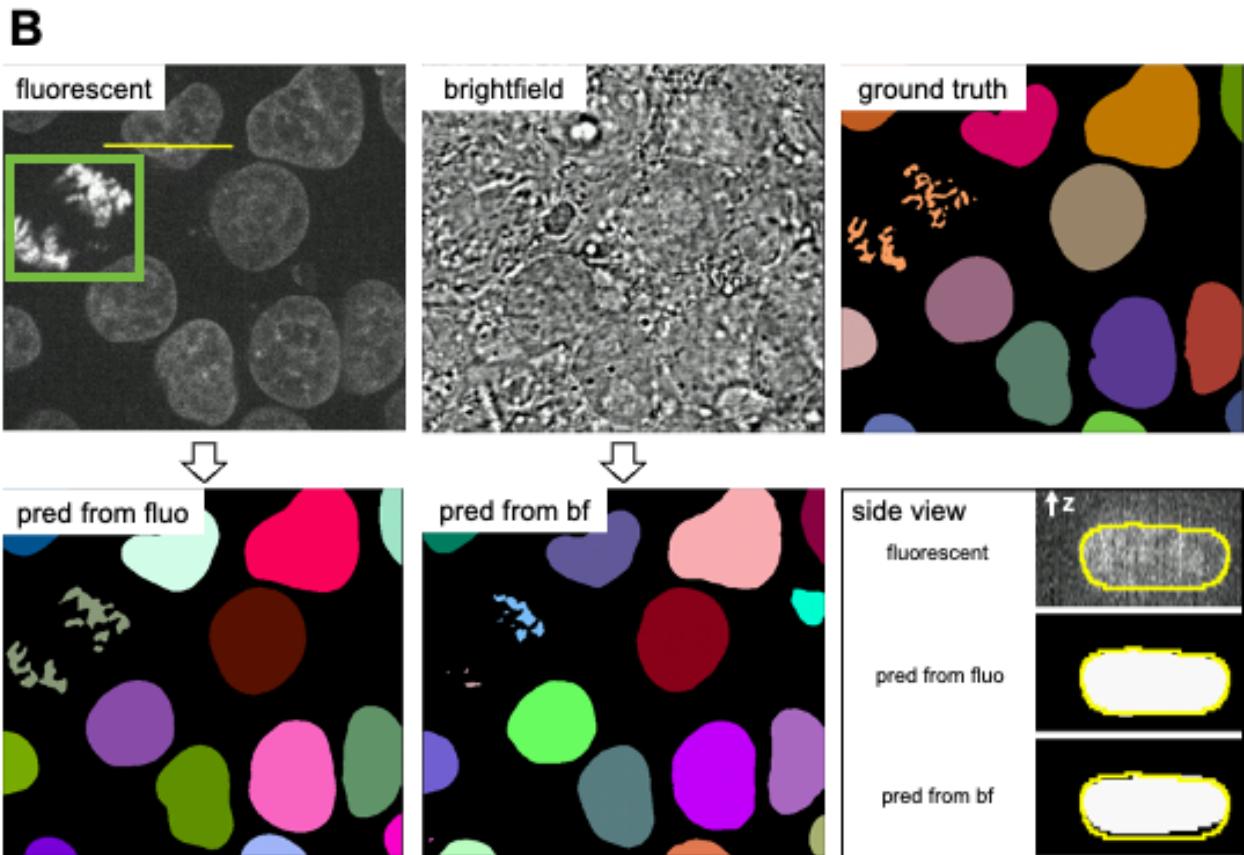
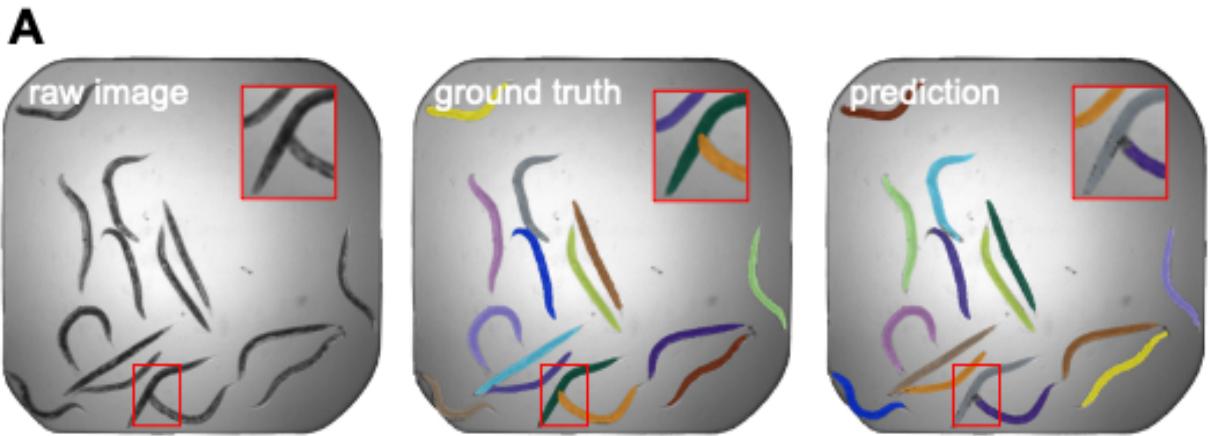


Figure 6: (A) Results 2D instance segmentation of *C. elegans*. Some minor errors can be observed in the zoom-in window, which could be refined with post-processing. (B) Results of 3D nuclear instance segmentation from fluorescent images and brightfield images. The green box in the fluorescent image highlights a mitotic example. The side view panel shows the segmentation of one specific nucleus along the line annotated in the fluorescent image from the side.

Generating synthetic microscopy images from binary Masks

Sometimes, being able to generate a large amount of synthetic microscopy images would be an important step for developing image analysis methods. For example, we can use the synthetic images to train other deep learning models. As long as the images can be synthesized good enough, we could have unlimited amount of training data “for free” for certain applications. In addition, the synthetic images could also be used to evaluate other models when validation data are not easy to obtain. We want to demonstrate that **MMV_Im2Im** can also be used to generate 2D/3D synthetic microscopy images with high realism, either in a supervised or unsupervised manner. To do this, we collect a subset data from the hiPS single cell image dataset [21]. For a 3D demonstration, we took the nuclear segmentation from the dataset as input, and the real H2B images as the training target. In the 2D case, we extract the middle Z-slice from NPM1 images as the training target, while using the NPM1

segmentation results as the input binary mask. With the paired “mask + microscopy image” data, we can train the model in a supervised fashion, or randomly shuffle the data to simulate the situation without paired data and train the model in an unsupervised fashion. Example results can be found in Figure 7.

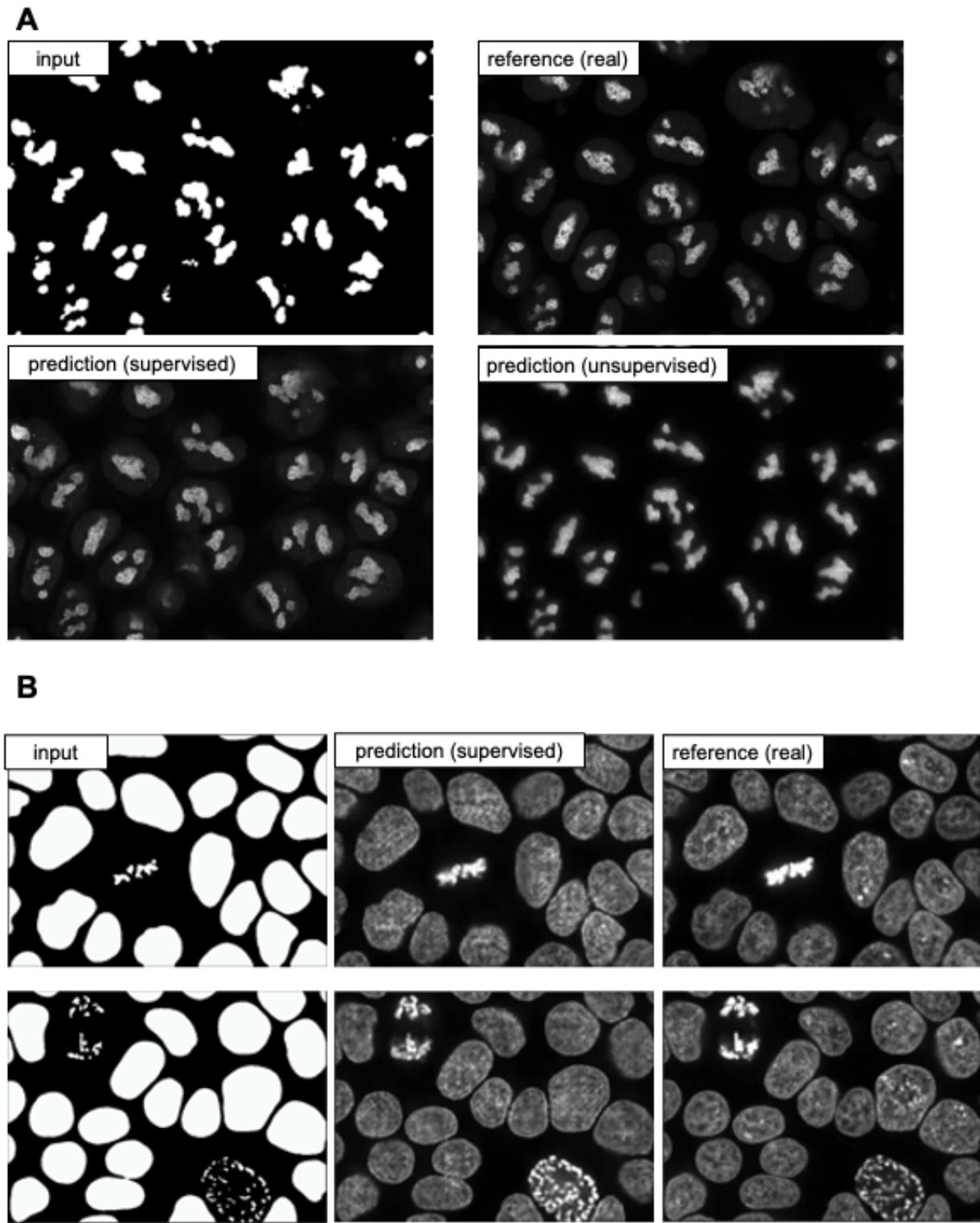


Figure 7: Overview of synthetic generated NPM1 fluorescence images.

Image denoising for microscopy images

MMV_Im2Im can also be used to significantly reduce image noise or restore the data from various sources of imaging artifacts, so as to increase the feasibility and efficiency in downstream analysis. In the current version of *MMV_Im2Im*, the restoration model can only be trained in a fully supervised

manner. Therefore, aligned low quality and high quality images are required for supervision, even though such pair data can be partially simulated [6]. Other methods, such as unsupervised learning based solutions [38], will be made available within **MMV_Im2Im** in future versions.

In this example, we present an image denoising demonstration with sample data from [39]. The goal is to increase the quality of low signal-to-noise ratio (SNR) images of nucleus-stained flatworm (*Schmidtea mediterranea*) and lightsheet images of *Tribolium castaneum* (red flour beetle) embryos. The models are trained with paired data acquired with low and high laser intensity on fixed samples, and then applied on live imaging data. For the nucleus-stained flatworm data (a test set of 20 images are available), the model achieved pearson correlation of 0.923 & 0.029 and structural similarity of 0.627 &pm 0.175. Based on the results in Figure 8, it can be observed that the low SNR images can be greatly improved. Systematic quantitative evaluations are necessary to confirm the biological validity, but beyond the scope of this paper.

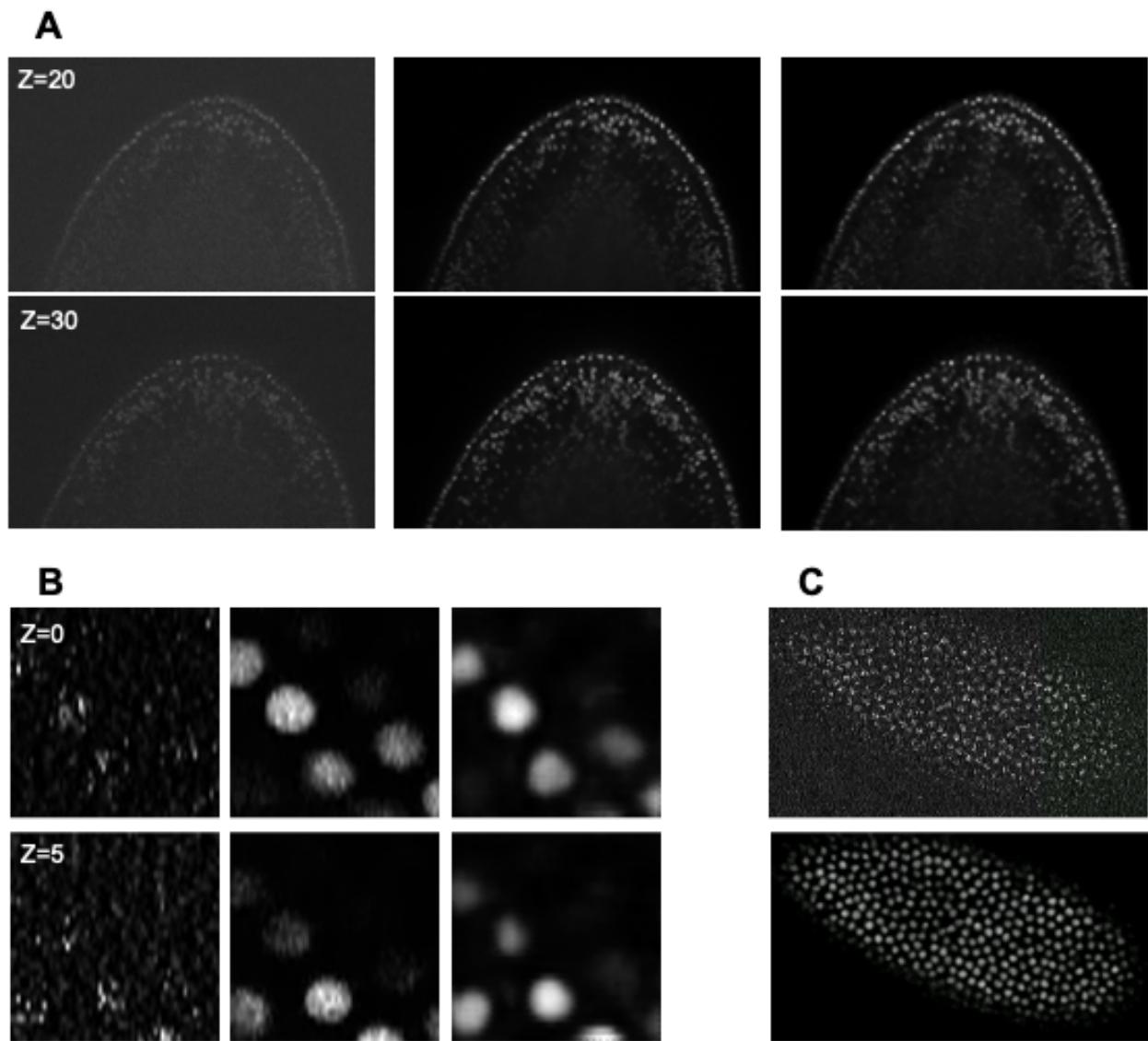
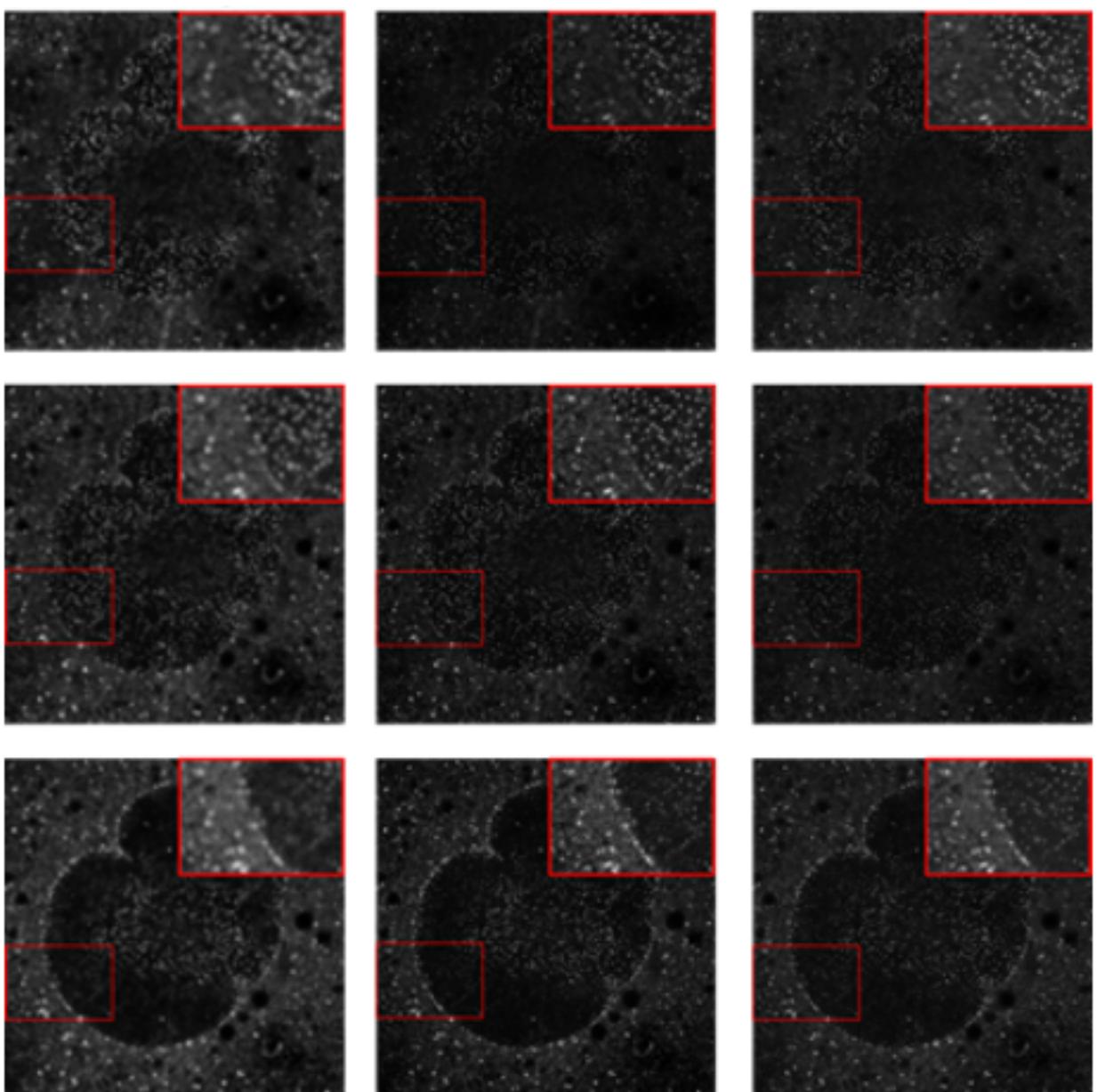
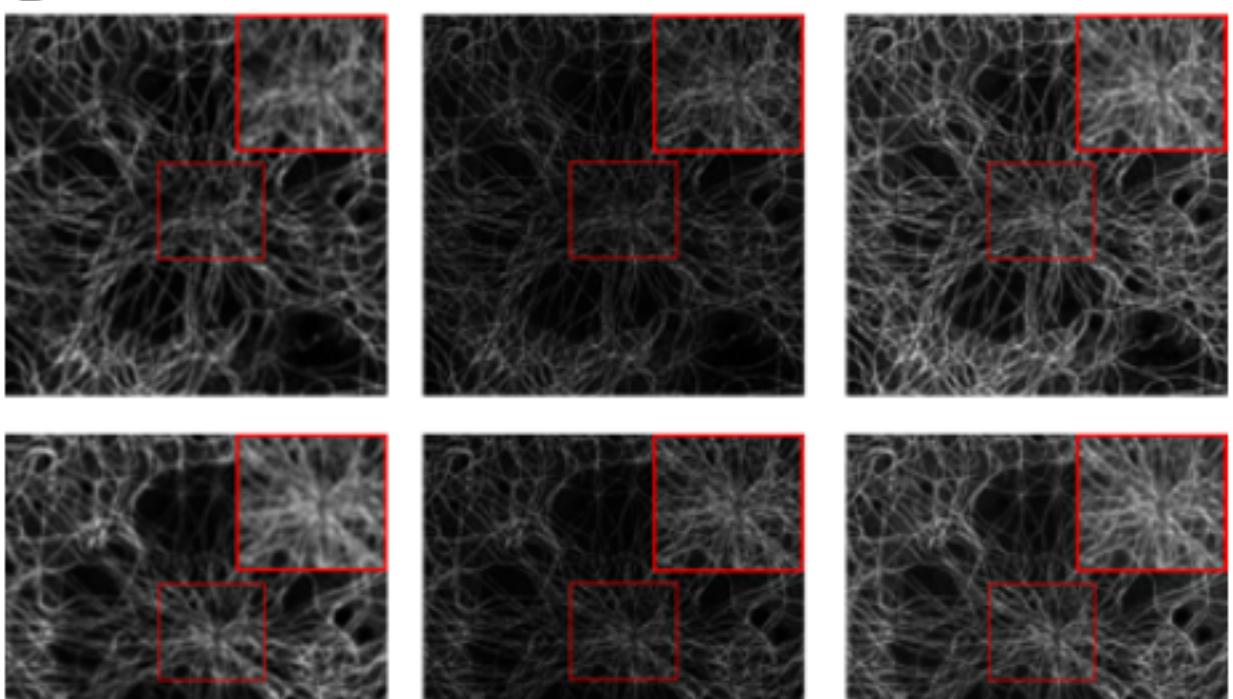


Figure 8: (A) Denoising results of 3D images of nucleus-stained flatworm at two different z-slices. Left: raw images (low SNR), middle: reference images (high SNR), right: predictions. (B) Denoising results of 3D lightsheet images of *Tribolium castaneum* (fixed samples) at two different z-slices. Left: raw images (low SNR), middle: Reference images (high SNR), right: predictions. (C) Denoising results of 3D lightsheet images of *Tribolium castaneum* (live samples) without high SNR reference. Top: the raw image, bottom: the prediction.

Imaging modality transformation from 3D confocal microscopy images to stimulated emission depletion (STED) microscopy images

Another important application of image-to-image transformation is imaging modality transformation [8], usually from one “cheaper” modality with lower resolution (e.g., larger field-of-view, easier to acquire and scale up) to another modality with higher resolution but expensive to obtain. Such models will permit a completely new way in assay development strategy to take advantage of all the benefits of the cheaper modality with lower resolution and still able to enhance the resolution computationally post hoc. To demonstrate the application of **MMV_Im2Im** in this scenario, we take an example dataset with paired 3D confocal and Stimulated Emission Depletion (STED) images of two different cellular structures [8]. Sample results are summarized in Figure 9. For microtubule, the model achieved pearson correlation of 0.779 ± 0.019 , while for nuclear pore complex, the pearson correlation is 0.784 ± 0.028 . Also, visual inspection can confirm the effectiveness of the models. Again, it is necessary to conduct further quantitative evaluation to ensure the validity in users’ specific problems.

A**B**

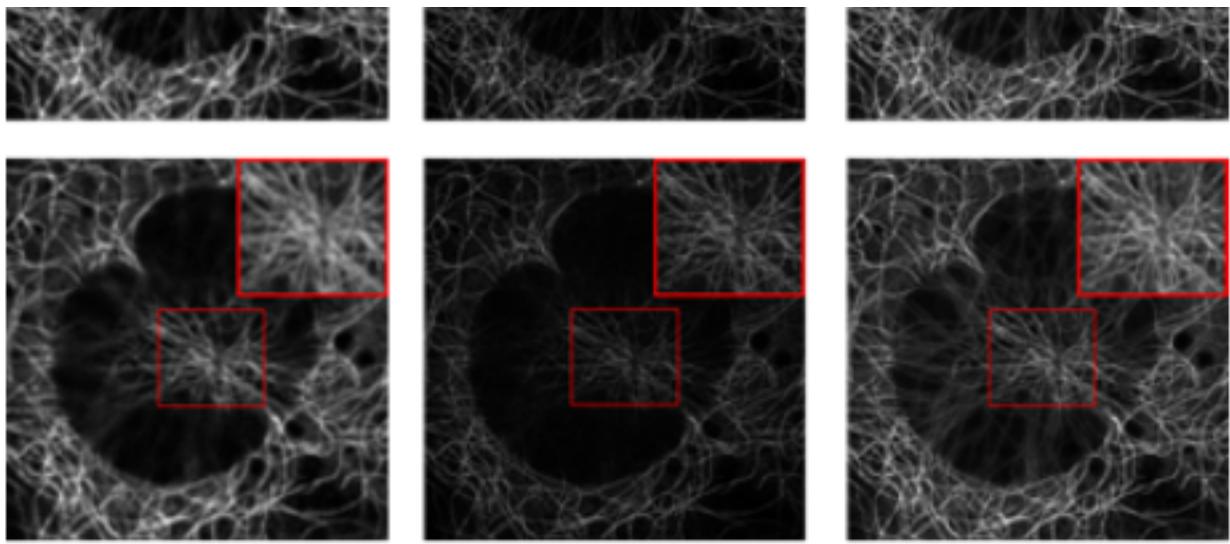


Figure 9: Example results of confocal-to-STED modality transformation of nuclear pore (A) and microtubule (B) in three consecutive z-slices. From left to right: raw confocal images, reference STED images, predicted images.

Staining transformation in multiplex experiments

Beyond common light microscope for bioimaging, deep learning based image-to-image transformation has become one of the important weapons in multiplex studies [5]. The capability of transforming tissue images among different stainings, such as immunohistochemistry (IHC) and multiplex immunofluorescence (mplF), would create new possibility for efficient co-registration, co-localization and quantitative analysis. In this section, we demonstrate the applicability of **MMV_Im2Im** in transforming tissue images from one staining to others with sample data from [5]. In this example, we train three different models to transform IHC images to images of standard hematoxylin stain, mplF nuclear (DAPI) and mplF LAP2beta (a nuclear envelope stain). Example results can be observed in Figure 10 to verify the results qualitatively.

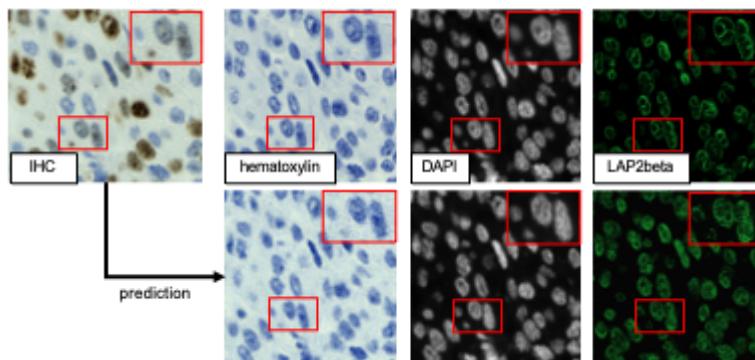


Figure 10: Qualitative visualization of staining transformation results with the **MMV_Im2Im** package.

Methods

Major components of the package

Discussions

This

References

1. **Enhanced Deep Residual Networks for Single Image Super-Resolution**
Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, Kyoung Mu Lee
2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (2017-07) <https://doi.org/gfxhp3>
DOI: [10.1109/cvprw.2017.151](https://doi.org/10.1109/cvprw.2017.151)
2. **Image-to-Image Translation with Conditional Adversarial Networks**
Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A Efros
2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017-07) <https://doi.org/gfrv9>
DOI: [10.1109/cvpr.2017.632](https://doi.org/10.1109/cvpr.2017.632)
3. **Panoptic Segmentation**
Alexander Kirillov, Kaiming He, Ross Girshick, Carsten Rother, Piotr Dollar
2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (2019-06) <https://doi.org/ggp9z4>
DOI: [10.1109/cvpr.2019.00963](https://doi.org/10.1109/cvpr.2019.00963)
4. **Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy**
Chawin Ounkomol, Sharmishtaa Seshamani, Mary M Maleckar, Forrest Collman, Gregory R Johnson
Nature Methods (2018-09-17) <https://doi.org/gd7d5f>
DOI: [10.1038/s41592-018-0111-2](https://doi.org/10.1038/s41592-018-0111-2) · PMID: [30224672](#) · PMCID: [PMC6212323](#)
5. **Deep learning-inferred multiplex immunofluorescence for immunohistochemical image quantification**
Parmida Ghahremani, Yanyun Li, Arie Kaufman, Rami Vanguri, Noah Greenwald, Michael Angelo, Travis J Hollmann, Saad Nadeem
Nature Machine Intelligence (2022-04) <https://doi.org/gqc7gd>
DOI: [10.1038/s42256-022-00471-x](https://doi.org/10.1038/s42256-022-00471-x)
6. **Deep learning-based point-scanning super-resolution imaging**
Linjing Fang, Fred Monroe, Sammy Weiser Novak, Lyndsey Kirk, Cara R Schiavon, Seungyoon B Yu, Tong Zhang, Melissa Wu, Kyle Kastner, Alaa Abdel Latif, ... Uri Manor
Nature Methods (2021-03-08) <https://doi.org/gjhgrw>
DOI: [10.1038/s41592-021-01080-z](https://doi.org/10.1038/s41592-021-01080-z) · PMID: [33686300](#) · PMCID: [PMC8035334](#)
7. **LIVECell—A large-scale dataset for label-free live cell segmentation**
Christoffer Edlund, Timothy R Jackson, Nabeel Khalid, Nicola Bevan, Timothy Dale, Andreas Dengel, Sheraz Ahmed, Johan Trygg, Rickard Sjögren
Nature Methods (2021-08-30) <https://doi.org/gmptqs>
DOI: [10.1038/s41592-021-01249-6](https://doi.org/10.1038/s41592-021-01249-6) · PMID: [34462594](#) · PMCID: [PMC8440198](#)
8. **Three-dimensional residual channel attention networks denoise and sharpen fluorescence microscopy image volumes**
Jiji Chen, Hideki Sasaki, Hoyin Lai, Yijun Su, Jiamin Liu, Yicong Wu, Alexander Zhovmer, Christian A Combs, Ivan Rey-Suarez, Hung-Yu Chang, ... Hari Shroff
Nature Methods (2021-05-31) <https://doi.org/gkbctn>
DOI: [10.1038/s41592-021-01155-x](https://doi.org/10.1038/s41592-021-01155-x) · PMID: [34059829](#)
9. **U-Net: Convolutional Networks for Biomedical Image Segmentation**

10. **Embedding-based Instance Segmentation in Microscopy**
Manan Lalit, Pavel Tomancak, Florian Jug
Proceedings of the Fourth Conference on Medical Imaging with Deep Learning
<https://proceedings.mlr.press/v143/lalit21a.html>
11. **Unsupervised data to content transformation with histogram-matching cycle-consistent generative adversarial networks**
Stephan J Ihle, Andreas M Reichmuth, Sophie Girardin, Hana Han, Flurin Stauffer, Anne Bonnin, Marco Stampanoni, Karthik Pattisapu, János Vörös, Csaba Forró
Nature Machine Intelligence (2019-09-16) <https://doi.org/ggwbv>
DOI: [10.1038/s42256-019-0096-2](https://doi.org/10.1038/s42256-019-0096-2)
12. **Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks**
Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A Efros
2017 IEEE International Conference on Computer Vision (ICCV) (2017-10) <https://doi.org/gfhw33>
DOI: [10.1109/iccv.2017.244](https://doi.org/10.1109/iccv.2017.244)
13. **PyTorchLightning/pytorch-lightning: 0.7.6 release**
William Falcon, Jirka Borovec, Adrian Wälchli, Nic Eggert, Justus Schock, Jeremy Jordan, Nicki Skafte, lr1dXD, Vadim Bereznyuk, Ethan Harris, ... Anton Bakhtin
Zenodo (2020-05-15) <https://doi.org/gqc7f9>
DOI: [10.5281/zenodo.3828935](https://doi.org/10.5281/zenodo.3828935)
14. **Averaging Weights Leads to Wider Optima and Better Generalization**
Pavel Izmailov, Dmitrii Podoprikin, Timur Garipov, Dmitry Vetrov, Andrew Gordon Wilson
34th Conference on Uncertainty in Artificial Intelligence 2018, UAI 2018
<http://auai.org/uai2018/proceedings/papers/313.pdf>
15. **ImageNet: A large-scale hierarchical image database**
Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, Li Fei-Fei
2009 IEEE Conference on Computer Vision and Pattern Recognition (2009-06)
<https://doi.org/cvc7xp>
DOI: [10.1109/cvpr.2009.5206848](https://doi.org/10.1109/cvpr.2009.5206848)
16. **Project MONAI**
The MONAI Consortium
Zenodo (2020-12-15) <https://doi.org/gqc7gb>
DOI: [10.5281/zenodo.4323059](https://doi.org/10.5281/zenodo.4323059)
17. **AllenCellModeling/aicsimageio: Types, Filesystem Management, and BioformatsReader Fixes**
Jackson Maxfield Brown, Jamie Sherman, Toloudis, Madison Swain-Bowden, Talley Lambert, Matte Bailey, Basu Chaudhuri, Gregory Johnson, Ian Hunt-Isaak, Nicholas-Schaub, ... Peter Sobolewski
Zenodo (2022-05-27) <https://doi.org/gqc7gc>
DOI: [10.5281/zenodo.6585658](https://doi.org/10.5281/zenodo.6585658)
18. **The Allen Cell and Structure Segmenter: a new open source toolkit for segmenting 3D intracellular structures in fluorescence microscopy images**
Jianxu Chen, Liya Ding, Matheus P Viana, HyeonWoo Lee, MFilip Sluezwski, Benjamin Morris, Melissa C Hendershott, Ruian Yang, Irina A Mueller, Susanne M Rafelski

19. **Open collaborative writing with Manubot**

Daniel S Himmelstein, Vincent Rubinetti, David R Slochower, Dongbo Hu, Venkat S Malladi, Casey S Greene, Anthony Gitter
PLOS Computational Biology (2019-06-24) <https://doi.org/c7np>
DOI: [10.1371/journal.pcbi.1007128](https://doi.org/10.1371/journal.pcbi.1007128) · PMID: [31233491](#) · PMCID: [PMC6611653](#)

20. **HeLa "Kyoto" cells under the scope**

Romain Guiet
Zenodo (2022-02-25) <https://doi.org/gqdkdm>
DOI: [10.5281/zenodo.6139958](https://doi.org/10.5281/zenodo.6139958)

21. **Robust integrated intracellular organization of the human iPS cell: where, how much, and how variable**

Matheus P Viana, Jianxu Chen, Theo A Knijnenburg, Ritvik Vasan, Calysta Yan, Joy E Arakaki, Matte Bailey, Ben Berry, Antoine Borensztein, Jackson M Brown, ...
Cold Spring Harbor Laboratory (2020-12-10) <https://doi.org/gqdkdk>
DOI: [10.1101/2020.12.08.415562](https://doi.org/10.1101/2020.12.08.415562)

22. **Left-Ventricle Quantification Using Residual U-Net**

Eric Kerfoot, James Clough, Ilkay Oksuz, Jack Lee, Andrew P King, Julia A Schnabel
Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges (2019) <https://doi.org/gqdkdp>
DOI: [10.1007/978-3-030-12029-0_40](https://doi.org/10.1007/978-3-030-12029-0_40)

23. **Attention U-Net: Learning Where to Look for the Pancreas**

Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Matthias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, ... Daniel Rueckert
Proceedings of Medical Imaging with Deep Learning 2018 <https://openreview.net/forum?id=Skft7cijM>

24. **Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images**

Ali Hatamizadeh, Vishwesh Nath, Yucheng Tang, Dong Yang, Holger R Roth, Daguang Xu
Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries (2022)
<https://doi.org/gqr93>
DOI: [10.1007/978-3-031-08999-2_22](https://doi.org/10.1007/978-3-031-08999-2_22)

25. **UNETR: Transformers for 3D Medical Image Segmentation**

Ali Hatamizadeh, Yucheng Tang, Vishwesh Nath, Dong Yang, Andriy Myronenko, Bennett Landman, Holger R Roth, Daguang Xu
2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) (2022-01)
<https://doi.org/gqr96>
DOI: [10.1109/wacv51458.2022.00181](https://doi.org/10.1109/wacv51458.2022.00181)

26. **High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs**

Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, Bryan Catanzaro
2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (2018-06)
<https://doi.org/ggv75b>
DOI: [10.1109/cvpr.2018.00917](https://doi.org/10.1109/cvpr.2018.00917)

27. **A Stochastic Polygons Model for Glandular Structures in Colon Histology Images**

Korsuk Sirinukunwattana, David RJ Snead, Nasir M Rajpoot

IEEE Transactions on Medical Imaging (2015-11) <https://doi.org/gqrg95>
DOI: [10.1109/tmi.2015.2433900](https://doi.org/10.1109/tmi.2015.2433900) · PMID: [25993703](https://pubmed.ncbi.nlm.nih.gov/25993703/)

28. **Gland segmentation in colon histology images: The glas challenge contest**
Korsuk Sirinukunwattana, Josien PW Pluim, Hao Chen, Xiaojuan Qi, Pheng-Ann Heng, Yun Bo Guo, Li Yang Wang, Bogdan J Matuszewski, Elia Bruni, Urko Sanchez, ... Nasir M Rajpoot
Medical Image Analysis (2017-01) <https://doi.org/c5t7>
DOI: [10.1016/j.media.2016.08.008](https://doi.org/10.1016/j.media.2016.08.008) · PMID: [27614792](https://pubmed.ncbi.nlm.nih.gov/27614792/)
29. **A method for normalizing histology slides for quantitative analysis**
Marc Macenko, Marc Niethammer, JS Marron, David Borland, John T Woosley, Xiaojun Guan, Charles Schmitt, Nancy E Thomas
2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro (2009-06)
<https://doi.org/bmbj4h>
DOI: [10.1109/isbi.2009.5193250](https://doi.org/10.1109/isbi.2009.5193250)
30. **Dense cellular segmentation for EM using 2D-3D neural network ensembles**
Matthew D Guay, Zeyad AS Emam, Adam B Anderson, Maria A Aronova, Irina D Pokrovskaya, Brian Storrie, Richard D Leapman
Scientific Reports (2021-01-28) <https://doi.org/gqdxxg>
DOI: [10.1038/s41598-021-81590-0](https://doi.org/10.1038/s41598-021-81590-0) · PMID: [33510185](https://pubmed.ncbi.nlm.nih.gov/33510185/) · PMCID: [PMC7844272](https://pubmed.ncbi.nlm.nih.gov/PMC7844272/)
31. **Cell Detection with Star-Convex Polygons**
Uwe Schmidt, Martin Weigert, Coleman Broaddus, Gene Myers
Medical Image Computing and Computer Assisted Intervention - MICCAI 2018 (2018)
<https://doi.org/ggnzqb>
DOI: [10.1007/978-3-030-00934-2_30](https://doi.org/10.1007/978-3-030-00934-2_30)
32. **Star-convex Polyhedra for 3D Object Detection and Segmentation in Microscopy**
Martin Weigert, Uwe Schmidt, Robert Haase, Ko Sugawara, Gene Myers
2020 IEEE Winter Conference on Applications of Computer Vision (WACV) (2020-03)
<https://doi.org/gjp4g9>
DOI: [10.1109/wacv45572.2020.9093435](https://doi.org/10.1109/wacv45572.2020.9093435)
33. **Cellpose: a generalist algorithm for cellular segmentation**
Carsen Stringer, Tim Wang, Michalis Michaelos, Marius Pachitariu
Nature Methods (2020-12-14) <https://doi.org/ghrgms>
DOI: [10.1038/s41592-020-01018-x](https://doi.org/10.1038/s41592-020-01018-x) · PMID: [33318659](https://pubmed.ncbi.nlm.nih.gov/33318659/)
34. **Splinedist: Automated Cell Segmentation With Spline Curves**
Soham Mandal, Virginie Uhlmann
2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI) (2021-04-13)
<https://doi.org/gqrg94>
DOI: [10.1109/isbi48211.2021.9433928](https://doi.org/10.1109/isbi48211.2021.9433928)
35. **Mask R-CNN**
Kaiming He, Georgia Gkioxari, Piotr Dollar, Ross Girshick
IEEE Transactions on Pattern Analysis and Machine Intelligence (2020-02-01)
<https://doi.org/gfxfwn>
DOI: [10.1109/tpami.2018.2844175](https://doi.org/10.1109/tpami.2018.2844175) · PMID: [29994331](https://pubmed.ncbi.nlm.nih.gov/29994331/)
36. **High-Throughput Screen for Novel Antimicrobials using a Whole Animal Infection Model**
Terence I Moy, Annie L Conery, Jonah Larkins-Ford, Gang Wu, Ralph Mazitschek, Gabriele Casadei, Kim Lewis, Anne E Carpenter, Frederick M Ausubel
ACS Chemical Biology (2009-06-29) <https://doi.org/bdwdfc>

37. **Fully Unsupervised Probabilistic Noise2Void**

Mangal Prakash, Manan Lalit, Pavel Tomancak, Alexander Krul, Florian Jug

2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI) (2020-04)

<https://doi.org/gjnw2v>

DOI: [10.1109/isbi45749.2020.9098612](https://doi.org/10.1109/isbi45749.2020.9098612)

38. **Interpretable Unsupervised Diversity Denoising and Artefact Removal**

Mangal Prakash, Mauricio Delbracio, Peyman Milanfar, Florian Jug

Proceedings of the Tenth International Conference on Learning Representations

<https://proceedings.mlr.press/v143/lalit21a.html>

39. **Content-aware image restoration: pushing the limits of fluorescence microscopy**

Martin Weigert, Uwe Schmidt, Tobias Boothe, Andreas Müller, Alexandr Dibrov, Akanksha Jain, Benjamin Wilhelm, Deborah Schmidt, Coleman Broaddus, Siân Culley, ... Eugene W Myers

Nature Methods (2018-11-26) <https://doi.org/gfkkfd>

DOI: [10.1038/s41592-018-0216-7](https://doi.org/10.1038/s41592-018-0216-7) · PMID: [30478326](https://pubmed.ncbi.nlm.nih.gov/30478326/)