

Enhancing high-content imaging for studying microtubule networks at large-scale

Hao-Chih Lee
Sarah T. Cherng
Riccardo Miotto
Joel T. Dudley

Institute for Next Generation Healthcare
Department of Genetics and Genomics Sciences
Icahn School of Medicine at Mount Sinai
New York, NY 10065, USA

HAO-CHIH.LEE@MSSM.EDU
SARAH.CHERNG@MSSM.EDU
RICCARDO.MIOTTO@MSSM.EDU
JOEL.DUDLEY@MSSM.EDU

Abstract

Given the crucial role of microtubules for cell survival, many researchers have found success using microtubule-targeting agents in the search for effective cancer therapeutics. Understanding microtubule responses to targeted interventions requires that the microtubule network within cells can be consistently observed across a large sample of images. However, fluorescence noise sources captured simultaneously with biological signals while using wide-field microscopes can obfuscate fine microtubule structures. Such requirements are particularly challenging for high-throughput imaging, where researchers must make decisions related to the trade-off between imaging quality and speed. Here, we propose a computational framework to enhance the quality of high-throughput imaging data to achieve fast speed and high quality simultaneously. Using CycleGAN, we learn an image model from low-throughput, high-resolution images to enhance features, such as microtubule networks in high-throughput low-resolution images. We show that CycleGAN is effective in identifying microtubules with 0.93+ AUC-ROC and that these results are robust to different kinds of image noise. We further apply CycleGAN to quantify the changes in microtubule density as a result of the application of drug compounds, and show that the quantified responses correspond well with known drug effects.

1. Introduction

Microtubules, composed of tubulin, are the major part of the cytoskeleton and are crucial for cell division, migration, trafficking, and signalling (Lodish et al., 2008). In particular, microtubules play a central role in the process of cell mitosis, which has resulted in microtubule-binding agents (MTAs) acting as effective anticancer agents by inducing mitotic arrest and cell death as cancer cells divide more rapidly than normal cells (Mukhtar et al., 2014). Continued study of microtubule responses to MTAs remain paramount due to significant discoveries related to difficulties in the use of MTAs to treat neoplastic cells, such as neuronal toxicity, and drug resistance (Orr et al., 2003).

High-throughput phenotypic screening is a powerful search tool for cancer drug candidates, which applies large-scale sets of compounds to cultured cells with the goal of quan-

tifying their effects by measuring changes in cellular phenotypes. High-content screening (HCS) is a type of high-throughput phenotypic screening that uses microscopy to image cellular components and outputs images as measured phenotypes. This is often conducted with fluorescence microscopy, where cellular components are labeled with fluorescent probes and visualized under a microscope. HCS has been applied to identify potential drug candidates for treating cerebral cavernous malformation (Gibson et al., 2015) and inhibiting bacterial growth (Sundaramurthy et al., 2013).

Unbiased HCS requires the ability to capture a large number of images of subcellular outcomes within a reasonable amount of time. For practical reasons, trade-off decisions are made between imaging speed and imaging resolution for HCS-specific experiments (Buchser et al., 2014). The consequences of such trade-offs are particularly challenging for HCS on microtubule networks for the purposes of drug discovery, as both image quality (i.e., recovering structures of microtubule networks) and speed (i.e., testing many compounds in sequence) are critical in the identification of effective drugs. The wide-field microscope, despite its superior speed, captures out-of-focus fluorescent signals that can greatly mask microtubules. Confocal microscopes are able to dramatically reduce fluorescent signals at the cost of slower image acquisition.

In this paper, we propose a computational solution to capitalize on the advantages of both imaging modalities to generate high quality images at sufficiently fast speeds. To computationally enhance the important structural signals of microtubule networks within HCS images, we propose using a generative adversarial networks model approach (Figure 1). We begin by training a model using high-resolution images of microtubule networks from low-throughput confocal images, collected from The Human Protein Atlas (Thul et al., 2017). Using the CycleGAN framework (Zhu et al., 2017), we demonstrate that the quality of HCS images can be enhanced by training a model to identify important features of microtubule networks from external reference images. With the removal of unwanted fluorescence-associated signals, the microtubule networks can be easily segmented from images using Otsu thresholding (Otsu, 1979). Finally, we apply this technique to verify whether expected drug effects on the density of microtubule networks can be observed from computationally enhanced HCS images.

Technical Significance The methodology presented here is an unsupervised approach to improve the quality of HCS images. The unsupervised nature of the proposed method is key to its applicability to HCS images, in which image noise may vary from batch to batch. The presence of such batch effects may lead to deteriorated performance of supervised models. We aim, specifically, to eliminate the need of supervised training using paired images, either taken simultaneously by another high-resolution imaging modality (Ouyang et al., 2018; Zhang et al., 2019; Wang et al., 2019) or generated by *ad hoc* simulation (Nehme et al., 2018; Weigert et al., 2018).

Clinical Relevance Our method offers new perspectives on the study of microtubule networks, a focal point of cancer drug discovery, by leveraging a fast but low-resolution imaging modality. This method can be generalized to other cellular components of interest with proper reference images.

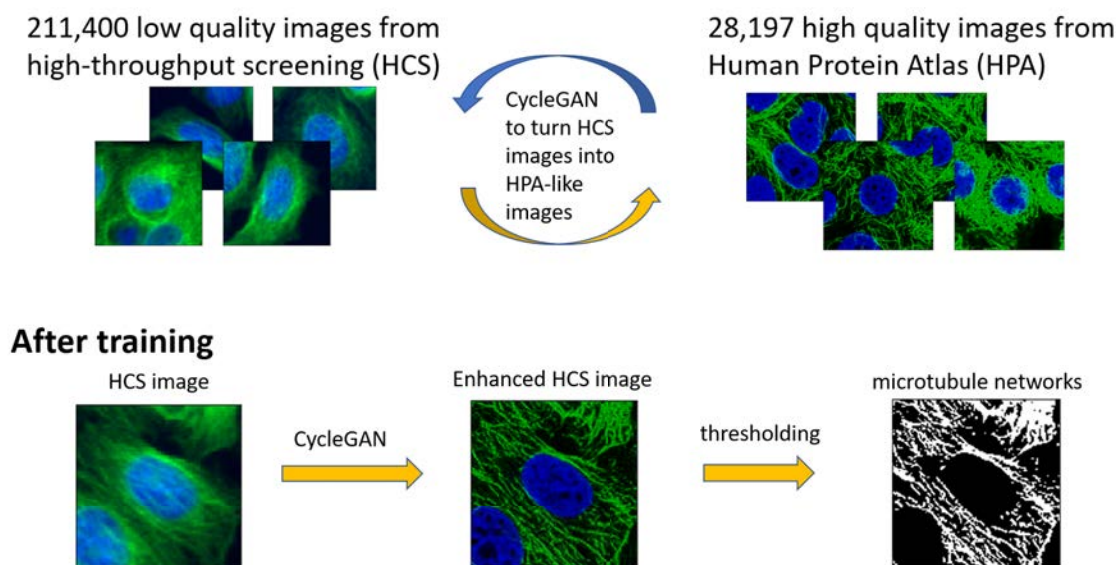


Figure 1: Illustration of a CycleGAN and its application to enhance the quality of HCS images.

2. Related Work

Various approaches have been developed to improve the quality and resolution of fluorescent microscopy. A super-resolution microscopy is the current state-of-the-art method, which breaks down the diffraction barrier using a specialized hardware design. For instance, simulation emission depletion microscopy uses a structured laser light to selectively deactivate fluorophores around a focal point to improve resolution (Hell and Wichmann, 1994). Photoactivated localization microscopy (PALM, Hess et al. (2006)) and stochastic optical reconstruction microscopy (STORM, Bates et al. (2007)) record time-lapsed images of blinking fluorophores to computationally localize proteins with nanometer accuracy. However, super-resolution microscopy is not a viable modality for high-throughput drug screening due to speed constraints.

Enhancing the resolution of natural images is long rooted in the field of computer vision (Yang et al., 2014; Nasrollahi and Moeslund, 2014). Recently, neural network models have been applied to enhance the resolution of natural photography (Dong et al., 2014; Ledig et al., 2017). This method was subsequently applied to fluorescent microscopy. For example, a supervised neural network can be trained to transform low resolution images into high resolution ones, either taken simultaneously by a different imaging modality, or generated by simulation. The content-aware image restoration (CARE) networks were trained on pairs of registered images of both high and low signal-to-noise ratios (SNR) to enable restoring biological structures from low SNR images (Weigert et al., 2018). Additionally, Wang et al. trained a convolutional neural network using images taken under different imaging modalities to improve image resolution. Deep STORM (Nehme et al., 2018) and ANNA-

PALM (Ouyang et al., 2018) both utilized paired low-resolution images of blinking emitters and their coordinates to increase speed for the reconstruction of super-resolved images.

Three strategies are generally used to quantify the structure of microtubules from images. The first approach utilizes direct measurements from image segmentation of microtubules. For example, Faulkner et al. applied local thresholding to segment microtubules and analyze the length of its morphological skeleton. Jacques et al. developed a line search algorithm to connect neighboring pixels with comparable values to detect microtubules. This approach relies on the successful identification of individual microtubules, which can be challenging for HCS imaging due to its susceptibility to low image quality (Boudaoud et al., 2014). The second approach attempts to quantify the local coherent pattern of microtubules via image textures, utilizing patterns extracted from image gradients (Karlou et al., 1999), nematic tensors (Boudaoud et al., 2014; Tsugawa et al., 2016) and Haralick textures (Berges et al., 2017). The third approach is the generative modeling approach that searches for parameters to simulate the most realistic images of microtubule networks (Shariff et al., 2010). Our work aims at improving the applicability of the first approach by enhancing the quality of HCS images.

3. Methods

3.1. CycleGAN for enhancing fluorescent images

Generative adversarial networks (GANs, Goodfellow et al. (2014)) are a class of neural network models that map samples from a prior distribution to the data distribution. GANs consist of two neural networks: a discriminator $D(x)$ and a generator $G(z)$. The generator performs the primary task of mapping a sample from the prior distribution p_z to a data point while the discriminator classifies how similar this generated data point is compared to the real data. Both neural networks are trained simultaneously by solving a min-max game, in which the generator seeks to create samples with with the goal of fooling the discriminator. The discriminator seeks to learn features of the generated data so that it is able to discriminate generated samples from real data points. The objective function is originally formulated as

$$\mathcal{L}_{GAN}(D, G) = \mathbb{E}_{x \sim data} \log(D(x)) + \mathbb{E}_{z \sim p_z} \log(1 - D(G(z))).$$

It has been shown that GANs can generate very realistic images, with greater clarity than ones generated by autoencoders (Goodfellow et al., 2014).

While GANs are designed as a random sampler from the data distribution, conditional GANs generate images based on input images (Isola et al., 2017). The generator of the conditional GAN learns a mapping from an input image x and a random noise vector z , to a generated image y , while the discriminator remains the same as in the original GAN. To enforce the consistency between generated images y and input images x , Isola et al. proposes the use of the L_1 discrepancy to penalize the model, i.e.,

$$\mathcal{L}_{L_1}(G) = \mathbb{E}_{x,y,z} \|G(x, z) - y\|_1,$$

which requires supervised pairs of trained images x and y . The conditional GAN is then trained to minimize the objective function

$$\mathcal{L}_{cGAN}(D, G) = \mathcal{L}_{GAN}(D, G) + \lambda \mathcal{L}_{L_1}(G)$$

In the context of image-to-image translation, this model is often referred to as the ‘‘pix2pix’’ model.

CycleGANs take conditional GANs one step further. The primary goal of a cycleGAN is to transform images from a source distribution to an image from a target distribution without paired training examples (Zhu et al., 2017). For example, CycleGAN, trained on paintings of an artist, can render natural photographs into the style of a given artist. CycleGAN consists of four neural networks: two generators $G_{A \rightarrow B}(x)$ and $G_{B \rightarrow A}(x)$ that map the source and target distributions from one to another and two discriminators $D_A(x)$ and $D_B(x)$ to judge how realistic the generated samples are. To ensure that the generators retain information of its domain images, Zhu et al. proposed using cycle-consistency loss:

$$\begin{aligned} \mathcal{L}_{\text{consistency}} = & \mathbb{E}_{A \sim \text{source images}} \|G_{B \rightarrow A}(G_{A \rightarrow B}(A)) - A\|_1 \\ & + \mathbb{E}_{B \sim \text{target images}} \|G_{A \rightarrow B}(G_{B \rightarrow A}(B)) - B\|_1, \end{aligned}$$

which measures only self-reconstruction errors and thus circumvents the need of supervised paired training samples. The CycleGAN is trained to minimize the objective function

$$\mathcal{L}_{obj} = \mathcal{L}_{GAN}(D_A, G_{B \rightarrow A}) + \mathcal{L}_{GAN}(D_B, G_{A \rightarrow B}) + \lambda \mathcal{L}_{\text{consistency}}.$$

One of the fundamental assumptions of CycleGANs is that there exists a relationship between the source and target domains. For example, two pictures with different styles are rendered differently, despite capturing the same object. Such an assumption would hold true for images of the same cell samples taken under different microscopes. Images from a microscope can be seen as the output of an image formation process which takes labeled biological structures as input and returns an array of pixel intensity. The image formation processes of two different imaging modalities (e.g., HCS and high-resolution confocal images) A and B can be modeled as $f_A(X)$ and $f_B(X)$ respectively, where X indicates a potential configuration of the biological structures of interest. Assuming we can reverse the image formation process f_B by a reconstruction algorithm, taking the composition of $f_A \circ f_B^{-1}$ could, in principle, transfer a image taken under one imaging modality B to an image taken under a different imaging modality A and vice-versa. The generators of the CycleGAN can thus approximate the composition functions $G_{A \rightarrow B} \approx f_B \circ f_A^{-1}$ and $G_{B \rightarrow A} \approx f_A \circ f_B^{-1}$ without solving the reconstruction inverse problems. Though this theoretically justifies the use of CycleGAN, we seek to design a prototype and evaluate its empirical performance on enhancing fluorescent imaging.

3.2. Data Collection

We use images from The Human Protein Atlas to curate a reference set of high resolution images. The Human Protein Atlas is a database that archives images of 12,073 protein taken from 64 cell lines (Thul et al., 2017). Among these images, the cell nucleus and microtubules are typically imaged at the same time as the reference coordinate for other

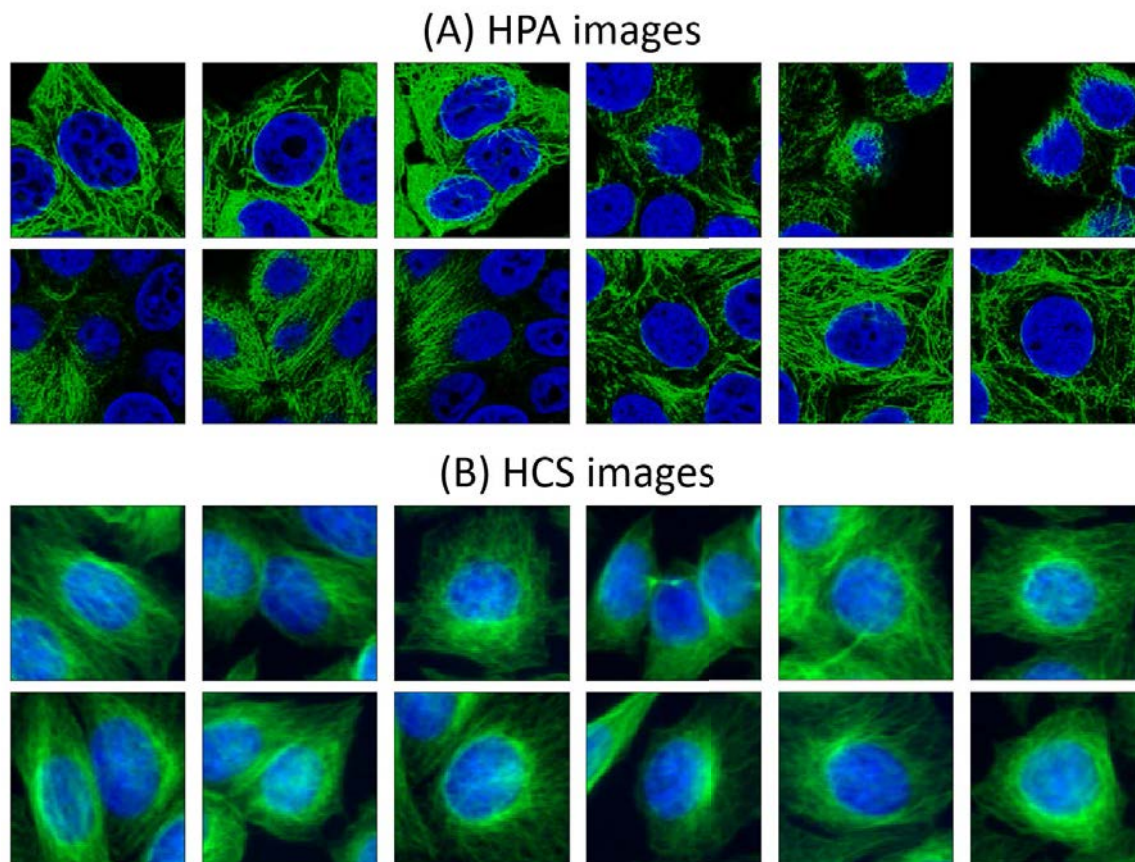


Figure 2: Examples of the (A) human protein atlas (HPA) and (B) high-content screening (HCS) image patches. Cell nuclei and microtubules are shown in blue and green, respectively.

proteins of interest. Images are collected under a $63\times$ confocal microscope, which reduces out-of-focus light. We select 1,836 images taken from the MCF7 cell line, which is the same cell line used in our HCS image data. Images are down-sampled by 4-fold and re-sized from 2048×2048 to 512×512 , from which we generate single-cell image patches. We first identify the cell nucleus using Otsu segmentation (Otsu, 1979). Image patches of size 128×128 , centered at the centroid of the cell nucleus, are then cropped. We only keep image patches that fall entirely within the original image. In total, we collect 28,197 single cell image patches. Each image patch contains a nucleus channel and a microtubule channel. Figure 2 shows examples of the high resolution image patches described here. We refer to this set of high-resolution image patches as the “HPA images” moving forward.

We also process the image set BBBC021v1 (Caie et al., 2010), available from the Broad Bioimage Benchmark Collection (Ljosa et al., 2012). BBBC021v1 is a HCS experiment on MCF7 cells. 13,200 images are collected using a 20x wide-field microscope. Each image contains a nucleus, microtubule, and actin filament channel. We retain only the nucleus and microtubule channels to match the settings of our HPA image set. Using the same

procedure to select images for the HPA set, we generate 84×84 images patches centered at the cell nucleus from untreated cells. In total, we collect 211,400 single-cell image patches from cells without the application of drug treatment for model training. We choose the size of image patches so that physical size of cell nucleus matches with those in our HPA images, as explained in the next section. We refer to this set of high-throughput, wide-field microscope images, as the ‘‘HCS images’’ set.

3.3. Implementation details

3.3.1. DATA AUGMENTATION

For both training and inference, we first re-size the low-resolution HCS images from 84×84 to 128×128 to ensure that the average size of cell nuclei in the HPA and HCS image sets are roughly equivalent. The magnification factor is set to $128/84 \approx 1.523$, which we determine by comparing the average size of cell nuclei in both HPA and HCS images. We apply Otsu thresholding to both image sets and determine that, on average, a nucleus contains 2,172 and 957 pixels for HPA and HCS images, respectively and therefore $128 \times \sqrt{957/2172} \approx 84$. We then randomly crop 128×128 images into 64×64 sized images when training.

3.3.2. PARAMETERS FOR TRAINING

We use a residual network with 9 repeated blocks for the generator and a 4-layer fully-connected convolutional neural network, sharing the same architecture as the 70×70 PatchGAN for the discriminator (Isola et al., 2017; Zhu et al., 2017). The regularization parameter λ is set to 10. We apply the ADAM optimizer (Kingma and Ba, 2014) for all experiments. We use a constant step size of 0.0002 for the first 20 epochs and then linearly reduce step size to 0 for the next 20 epochs. A batch of 8 samples is used for each iteration.

4. Results and Discussion

4.1. Evaluation by simulation

We conduct 3 simulations to evaluate the performance of the CycleGAN model. The source code is available at <https://github.com/DudleyLab/widefield2confocal>.

1. **Case 1 (blurred images):** We randomly split HPA images into 45/45/10 training, validation and testing sets. The training set contains 12,644 unprocessed, high-quality images while the validation and testing set contain 12,661 and 2892 degraded images, respectively. Using training and validation sets for model training, our goal is to evaluate the quality of reconstructed images from the testing set. The train/validate/test split is conducted on raw images that are downloaded from the Human Protein Atlas, but not cropped image patches, to avoid overlapping regions that may occur in the three sets of images. The 12,661 image patches in the validation set and 2,892 image patches in the testing set are down-sampled from 128×128 to 84×84 image patches, and blurred with a Gaussian kernel (bandwidth = 1), with the addition of Gaussian noise (standard deviation = 3). Specifically, we simulate a degraded image $I_{simulated}$ by

$$I_{simulated} = g_{\sigma=1} * I_{tubule} + 3 \cdot \epsilon$$

where $g_{\sigma=1}$ is the Gaussian kernel with a bandwidth of 1, ϵ is the normal Gaussian noise, and $*$ is the convolution.

2. **Case 2 (bleed-through signals)**: We keep the same train/validate/test split as in case 1. In addition to down-sampling, blurring, and the addition of noise, we introduce a bleed-through signal to the microtubule channel from the nucleus channel (Buchser et al., 2014). Specifically, we simulate an image $I_{simulated}$ by

$$I_{simulated} = g_{\sigma=1} * I_{tubule} + p \cdot (g_{\sigma=5} * I_{nucleus}) + 3 \cdot \epsilon$$

where $g_{\sigma=1}$ and $g_{\sigma=5}$ are Gaussian kernels with bandwidths of 1 and 5, respectively, and ϵ is normal Gaussian noise. We use $p = 0.2, 0.5$, and a random number ranging from 0.2 to 0.5 for cases 2a, 2b, and 2c respectively.

3. **Case 3 (diffusive fluorescent probes)**: We simulate the impact of non-specific signal from diffusive dyes (Toseland, 2013). The simulated images are generated as

$$I_{simulated} = (1 - p)(g_{\sigma=1} * I_{tubule}) + p \cdot (g_{\sigma=5} * I_{tubule}) + 3 \cdot \epsilon.$$

The first term is the true microtubule signal convoluted with the point spread function. The second term simulates the non-specific signal from dyes diffusing away from the microtubules. We simulate this effect by convolving true microtubule signals with a larger Gaussian kernel. We use the weighted average of two simulated signals to ensure that the pixel values remain in the same range. We use $p = 0.2, 0.5$, and a random number ranging from 0.2 to 0.5 for cases 3a, 3b, and 3c respectively.

All simulated images are binned as 8-bit integers.

Using the training and validation sets, we train a CycleGAN model for each simulation separately. After training, each model is applied to the corresponding testing images to recover microtubule signals. For comparison, we use the pix2pix model and the Richardson-Lucy deconvolution algorithm as baseline methods. For the pix2pix model, we create a set of simulated images from training images using the same setting as in case 1. A pix2pix model is then trained using only the training images and the corresponding simulation images. This pix2pix model is then applied to other testing image sets to recover microtubule signals. We also apply the Richardson-Lucy deconvolution algorithm to test images using $g_{\sigma=1}$ as the point spread function. For Richardson-Lucy deconvolution, we use scikit-image implementation (Van der Walt et al. (2014)). Figure 3 and 4 present examples of simulated images and reconstructed images using methods that we describe in case 2c and 3c.

We use three metrics to measure performance, with original test images acting as ground-truth.

1. Structural similarity index (SSIM, Wang et al. (2004)) and the multi-scale structural similarity index (MSSSIM, Wang et al. (2003)). SSIM and MSSSIM are scores designed to measure the perception-based similarity of two images in terms of luminance, contrast, and structure, but not the absolute difference in pixel values. Both scores are commonly used to benchmark the quality of images generated by GANs (Kurach et al., 2018). SSIM and MSSSIM both range from -1 and 1. A higher score indicates greater similarity between two images.

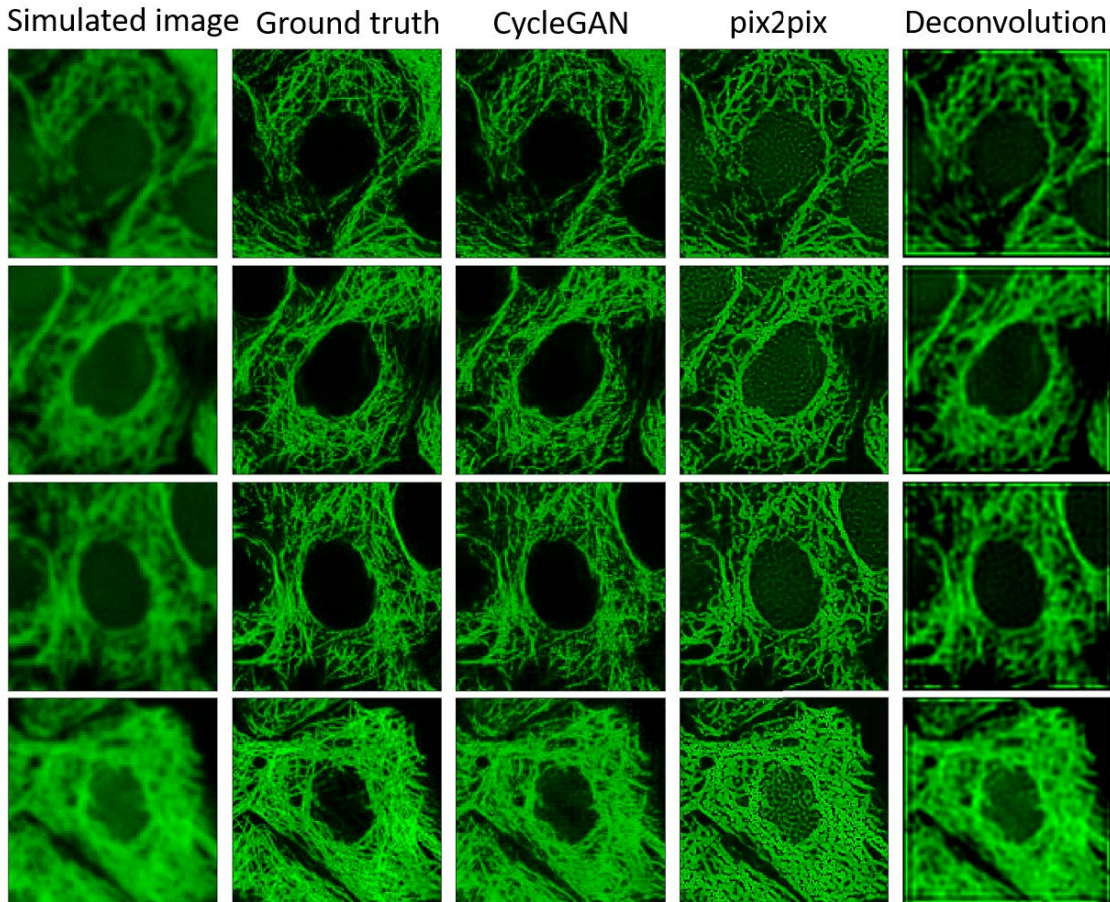


Figure 3: Examples of reconstructed images in case 2c.

2. We also evaluate whether the intensity of reconstructed signals indicate the true existence of microtubules. We label pixels in an image with values larger than the average as positive. The pixel value of the reconstructed image is then used as a logit to predict positive labels. Finally, the area under ROC-curve (AUC-ROC) is computed as a metric to measure prediction accuracy.

The results our case studies are shown in table 1. In the first case (case 1), the pix2pix model outperforms the other two. However, in the presence of bleed-through signals (case 2) and diffusive fluorescent probes (case 3), the CycleGAN performs best. This suggests that the pixel2pixel model and Richardson-Lucy deconvolution, which both assume that blurring is due only to noise, may fail to handle noise sources that deviate from the designed image formation model (Figure 3 and Figure 4). On the contrary, CycleGAN can robustly adapt to three types of noise sources. In all three cases, the 0.93+ AUC-ROC achieved by our CycleGAN model suggests that recovered pixel values are informative to the presence of microtubules.

Figure 5 presents the performance of our CycleGAN during training iterations in cases 2 and 3. In all simulated scenarios, CycleGAN achieves high accuracy in the first 4 epochs.

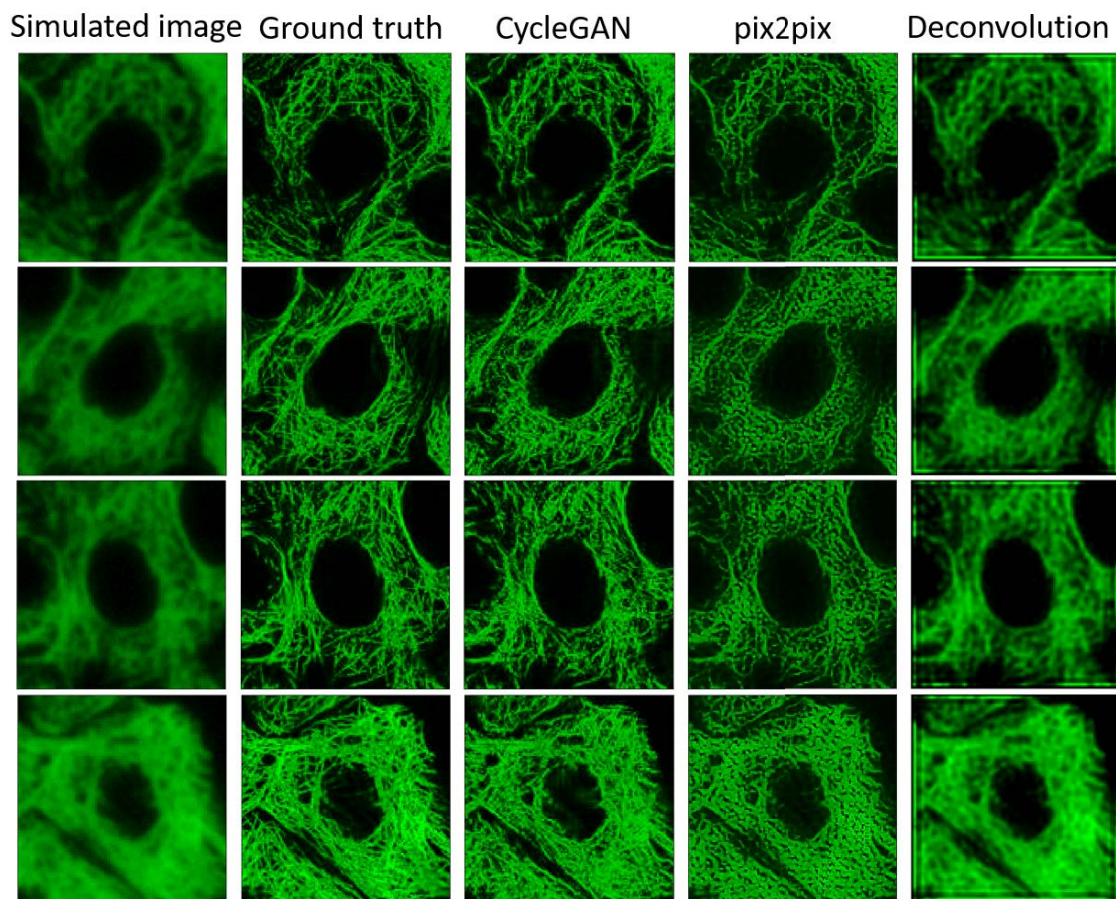


Figure 4: Examples of reconstructed images in case 3c.

Accuracy then either plateaus or improves gradually moving forward. We do not observe over-fitting within 40 epochs.

4.2. Application to high-throughput screening

Next, we apply our CycleGAN model to HCS images. We train a CycleGAN using 211,400 HCS image patches of untreated cells as the source images, and 28,197 HPA image patches as the target images. Figure 6 presents examples of enhanced images. We note two qualitative observations:

1. CycleGAN seems to adjust uneven illumination. Microtubule networks are generally denser around the cell nucleus than on the periphery. As a result, the fluorescent signals are generally weaker on the periphery in HCS images. CycleGAN is able to adjust this uneven illumination.
2. However, CycleGAN is also prone to eliminate microtubule signals within the region of a cell nucleus. This seems to be a biased correction resulting from using HPA images as a prior model since microtubules are also rarely present within the region of a cell nucleus in HPA images.

		CycleGAN	pix2pix	deconvolution
Case 1 blurred images	SSIM	65.8 ± 5.4	70.6 ± 5.6	65.5 ± 4.9
	MSSSIM	74.9 ± 4.5	77.1 ± 4.4	72.0 ± 4.8
	AUC-ROC	94.1 ± 2.2	94.8 ± 1.9	93.6 ± 2.1
Case 2a bleed-through signals, $p=0.2$	SSIM	62.2 ± 5.3	52.5 ± 7.2	50.8 ± 5.9
	MSSSIM	73.8 ± 4.5	64.5 ± 6.3	65.2 ± 4.7
	AUC-ROC	93.6 ± 4.8	89.6 ± 6.6	90.6 ± 5.9
Case 2b bleed-through signals, $p=0.5$	SSIM	61.4 ± 4.8	47.0 ± 8.7	47.0 ± 6.8
	MSSSIM	73.3 ± 4.0	57.5 ± 8.7	60.9 ± 6.1
	AUC-ROC	93.5 ± 3.6	83.8 ± 9.1	84.8 ± 9.2
Case 2c bleed-through signals, random p	SSIM	61.6 ± 5.5	50.0 ± 8.3	49.2 ± 6.5
	MSSSIM	73.2 ± 4.9	61.3 ± 8.1	63.5 ± 5.5
	AUC-ROC	93.4 ± 5.0	87.1 ± 8.3	88.4 ± 7.8
Case 3a diffusive fluorescent probes, $p=0.2$	SSIM	63.5 ± 5.7	61.1 ± 6.2	55.3 ± 5.3
	MSSSIM	73.0 ± 4.9	71.5 ± 5.1	64.8 ± 5.4
	AUC-ROC	93.8 ± 2.3	93.2 ± 2.3	92.6 ± 2.3
Case 3b diffusive fluorescent probes, $p=0.5$	SSIM	59.7 ± 5.8	46.5 ± 7.8	43.4 ± 6.2
	MSSSIM	70.0 ± 5.1	59.1 ± 6.5	54.6 ± 6.4
	AUC-ROC	93.0 ± 2.7	90.0 ± 3.2	91.1 ± 2.7
Case 3c diffusive fluorescent probes, random p	SSIM	61.0 ± 7.2	55.4 ± 8.6	50.6 ± 7.1
	MSSSIM	71.2 ± 5.9	66.9 ± 7.1	60.9 ± 6.8
	AUC-ROC	93.3 ± 2.6	92.0 ± 2.9	92.1 ± 2.5

Table 1: Performance of image reconstruction.

Based on the first observation of CycleGAN correcting uneven illumination, it is possible to segment microtubules from enhanced images using Otsu thresholding. Figure 7 shows examples of segmented microtubule networks. As a baseline, we apply the Sobel edge detector, followed by Otsu thresholding, to the original HCS images. Images that are reconstructed by both Richardson-Lucy deconvolution and pix2pix models, which are both trained using the first simulation case, are also segmented by Otsu thresholding for comparison. While there are practical difficulties in quantifying the accuracy of segmenting microtubule networks in HCS images, we note that the AUC-ROC metric described in section 4.1 evaluates the performance of this task using simulated data.

Lastly, we demonstrate an application of the proposed method to quantify drug effects on microtubule networks. We first apply CycleGAN to images of drug-treated cells and obtain segmented microtubule networks using Otsu segmentation. We then computed the number of segmented pixels that are no more than 8 pixels away from nuclei. The microtubule density is then defined as the ratio of the number of segmented pixels to the total area of this perinuclear region, excluding the region of cell nucleus itself. We restrict measurement within the perinuclear regions to control the cell size as a variable that can confound the estimation of microtubule density. 200 cells are randomly selected for each drug-concentration combination to estimate the drug effect on microtubule density.

Figure 8 shows the average microtubule density from cells treated with drugs of 3 known mechanisms of action. The dashed line represents the average microtubule density computed

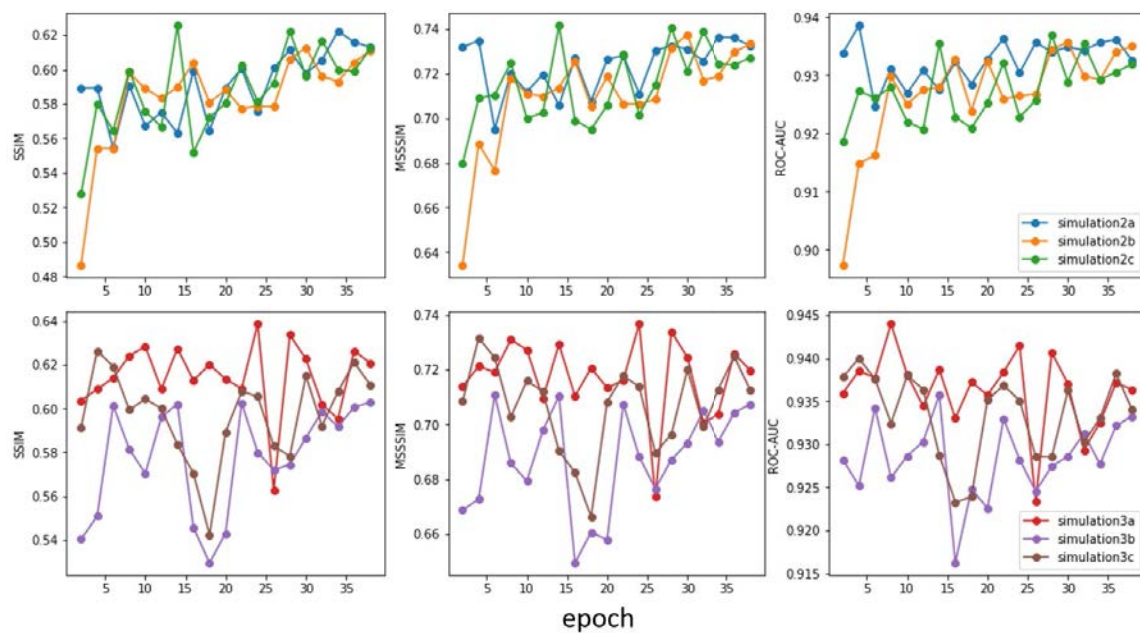


Figure 5: Performance during training iterations.

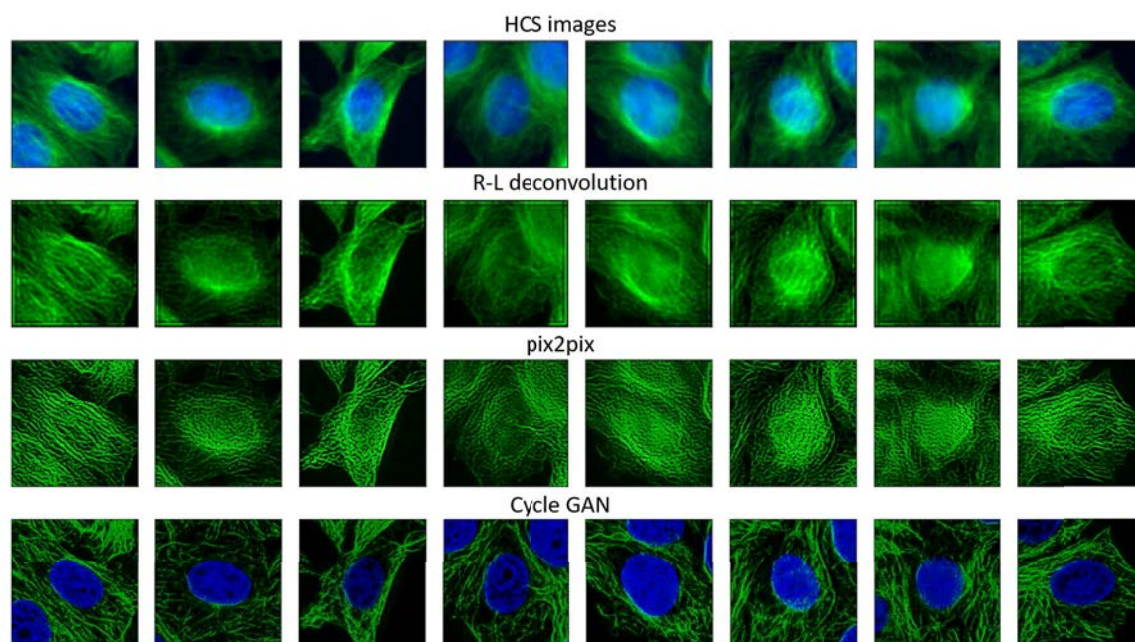


Figure 6: Examples of reconstructed HCS images. The Richardson-Lucy deconvolution and pix2pix model are applied only to the microtubule channel while the CycleGAN is trained using both the microtubule and nucleus channel. Cell nuclei and microtubules are shown in blue and green respectively.

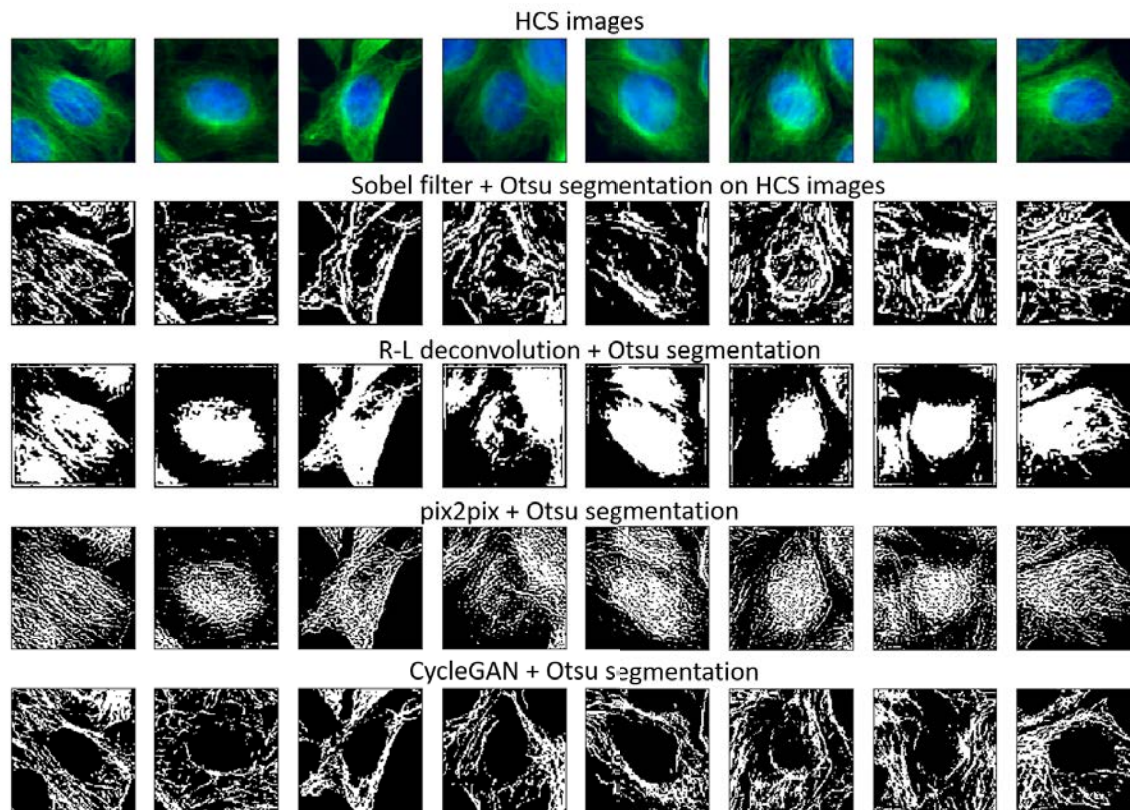


Figure 7: Examples of segmented microtubule networks.

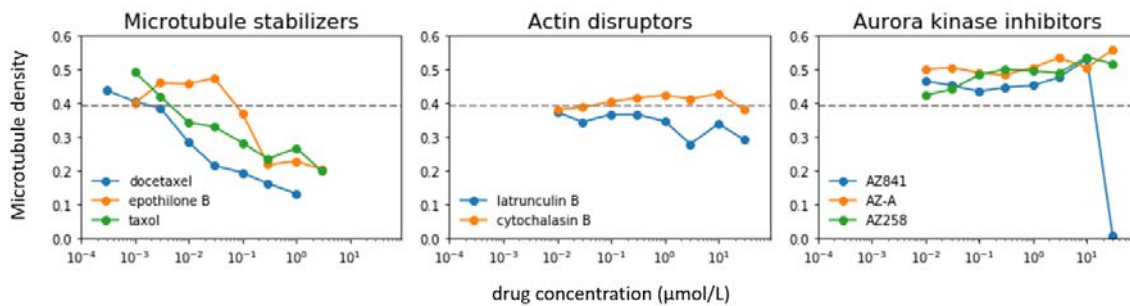


Figure 8: The dose-response curves of the effects of 8 drugs on microtubule density.

from 1,000 untreated cells. We observe that microtubule stabilizers, which are chemicals that interfere the growth and shrinkage of microtubules (Yvon et al., 1999; Jordan and Wilson, 2004) and promote microtubule polymerization (Jordan et al., 1993), induce higher microtubule density at low dosages, followed by a steep decline at high dosages. We note that the sharp drop to low microtubule density at high drug concentration levels is due to the increasing number of dying cells with shrunken cytoplasm (supplementary figure 1). We observe minor effects on microtubules for actin disruptors (supplementary figure 2). Lastly, we observe elevated microtubule density on cells treated with the Aurora kinase inhibitor, as Aurora kinase inhibitors seem to interrupt mitotic progression that leads to multinucleation (supplementary figure 3).

5. Conclusions and Future Work

In this paper, we present a framework to enhance the quality of HCS images using CycleGAN. Our empirical results demonstrate that CycleGAN can learn an unsupervised image model from high-quality images to identify relevant signals in low-quality images. In particular, we also show that CycleGAN can robustly adjust different types of image noise to achieve accurate identification of microtubules. Despite its superior performance, we observe biased removal of microtubules within the region of the cell nucleus, a bias likely inherited from curated reference images that we plan to resolve in future work.

In applying our method to HCS data, it is possible to identify hundreds of thousands of segmented microtubule networks from drug-treated cells. Here, we specify a case study to demonstrate the potential of this method by first analyzing the microtubule density. We plan to incorporate feature extraction techniques such as Hough transformation, to further analyze the lengths, orientations, and numbers of microtubules under different drug treatments. This new resource will allow us to discover new insights into the essential role of microtubule networks in cancer drug discovery research. Our method offers an opportunity for large-scale mining of microtubule networks. Importantly, this method is not limited to the study of microtubules alone. Indeed, with appropriate low-throughput, high resolution training images, the methods presented here may also effectively sharpen images of other subcellular components of interest.

Acknowledgments

The authors would like to thank support from the Hasso Plattner Foundation and a courtesy GPU donation from NVIDIA.

References

- Mark Bates, Bo Huang, Graham T Dempsey, and Xiaowei Zhuang. Multicolor super-resolution imaging with photo-switchable fluorescent probes. *Science*, 317(5845):1749–1753, 2007.
- Nina Berges, Katharina Arens, Verena Kreuzsch, Rainer Fischer, and Stefano Di Fiore. Toward discovery of novel microtubule targeting agents: A snap-tag-based high-content

- screening assay for the analysis of microtubule dynamics and cell cycle progression. *SLAS DISCOVERY: Advancing Life Sciences R&D*, 22(4):387–398, 2017.
- Arezki Boudaoud, Agata Burian, Dorota Borowska-Wykret, Magalie Uyttewaal, Roman Wrzalik, Dorota Kwiatkowska, and Olivier Hamant. Fibriltool, an imagej plug-in to quantify fibrillar structures in raw microscopy images. *Nature protocols*, 9(2):457, 2014.
- William Buchser, Mark Collins, Tina Garyantes, Rajarshi Guha, Steven Haney, Vance Lemmon, Zhuyin Li, and O Joseph Trask. Assay development guidelines for image-based high content screening, high content analysis and high content imaging. In *Assay Guidance Manual*. Eli Lilly & Company and the National Center for Advancing Translational Sciences, 2014.
- Peter D Caie, Rebecca E Walls, Alexandra Ingleston-Orme, Sandeep Daya, Tom Houslay, Rob Eagle, Mark E Roberts, and Neil O Carragher. High-content phenotypic profiling of drug response signatures across distinct cancer cells. *Molecular cancer therapeutics*, 9(6):1913–1926, 2010.
- Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Learning a deep convolutional network for image super-resolution. In *European conference on computer vision*, pages 184–199. Springer, 2014.
- Christine Faulkner, Ji Zhou, Alexandre Evrard, Gildas Bourdais, Dan MacLean, Heidrun Häweker, Peter Eckes, and Silke Robatzek. An automated quantitative image analysis tool for the identification of microtubule patterns in plants. *Traffic*, 18(10):683–693, 2017.
- Christopher C Gibson, Weiquan Zhu, Chadwick T Davis, Jay A Bowman-Kirigin, Aubrey C Chan, Jing Ling, Ashley E Walker, Luca Goitre, Simona Delle Monache, Savio Francesco Retta, et al. Strategy for identifying repurposed drugs for the treatment of cerebral cavernous malformation. *Circulation*, 131(3):289–299, 2015.
- Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.
- Stefan W Hell and Jan Wichmann. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics letters*, 19(11):780–782, 1994.
- Samuel T Hess, Thanu PK Girirajan, and Michael D Mason. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical journal*, 91(11):4258–4272, 2006.
- Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.
- Eveline Jacques, Jan Buytaert, Darren M Wells, Michal Lewandowski, Malcolm J Bennett, Joris Dirckx, Jean-Pierre Verbelen, and Kris Vissenberg. Microfilament analyzer, an

- image analysis tool for quantifying fibrillar orientation, reveals changes in microtubule organization during gravitropism. *The Plant Journal*, 74(6):1045–1058, 2013.
- Mary Ann Jordan and Leslie Wilson. Microtubules as a target for anticancer drugs. *Nature Reviews Cancer*, 4(4):253, 2004.
- Mary Ann Jordan, ROBERT J Toso, Doug Thrower, and Leslie Wilson. Mechanism of mitotic block and inhibition of cell proliferation by taxol at low concentrations. *Proceedings of the National Academy of Sciences*, 90(20):9552–9556, 1993.
- William J Karlon, Pin-Pin Hsu, Song Li, Shu Chien, Andrew D McCulloch, and Jeffrey H Omens. Measurement of orientation and distribution of cellular alignment and cytoskeletal organization. *Annals of biomedical engineering*, 27(6):712–720, 1999.
- Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- Karol Kurach, Mario Lucic, Xiaohua Zhai, Marcin Michalski, and Sylvain Gelly. The gan landscape: Losses, architectures, regularization, and normalization. *arXiv preprint arXiv:1807.04720*, 2018.
- Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017.
- Vebjorn Ljosa, Katherine L Sokolnicki, and Anne E Carpenter. Annotated high-throughput microscopy image sets for validation. *Nature methods*, 9(7):637–637, 2012.
- Harvey Lodish, James E Darnell, Arnold Berk, Chris A Kaiser, Monty Krieger, Matthew P Scott, Anthony Bretscher, Hidde Ploegh, Paul Matsudaira, et al. *Molecular cell biology*. Macmillan, 2008.
- Eiman Mukhtar, Vaqar Mustafa Adhami, and Hasan Mukhtar. Targeting microtubules by natural agents for cancer therapy. *Molecular Cancer Therapeutics*, 13(2):275284, 2014.
- Kamal Nasrollahi and Thomas B Moeslund. Super-resolution: a comprehensive survey. *Machine vision and applications*, 25(6):1423–1468, 2014.
- Elias Nehme, Lucien E Weiss, Tomer Michaeli, and Yoav Shechtman. Deep-storm: super-resolution single-molecule microscopy by deep learning. *Optica*, 5(4):458–464, 2018.
- George A Orr, Pascal Verdier-Pinard, Hayley McDaid, and Susan Band Horwitz. Mechanisms of taxol resistance related to microtubules. *Oncogene*, 22(47):7280–7295, 2003.
- Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66, 1979.

- Wei Ouyang, Andrey Aristov, Mickaël Lelek, Xian Hao, and Christophe Zimmer. Deep learning massively accelerates super-resolution localization microscopy. *Nature biotechnology*, 2018.
- Aabid Shariff, Robert F Murphy, and Gustavo K Rohde. A generative model of microtubule distributions, and indirect estimation of its parameters from fluorescence microscopy images. *Cytometry Part A: The Journal of the International Society for Advancement of Cytometry*, 77(5):457–466, 2010.
- Varadharajan Sundaramurthy, Rico Barsacchi, Nikolay Samusik, Giovanni Marsico, Jerome Gilleron, Inna Kalaidzidis, Felix Meyenhofer, Marc Bickle, Yannis Kalaidzidis, and Marino Zerial. Integration of chemical and rnaI multiparametric profiles identifies triggers of intracellular mycobacterial killing. *Cell host & microbe*, 13(2):129–142, 2013.
- Peter J Thul, Lovisa Åkesson, Mikaela Wiking, Diana Mahdessian, Aikaterini Geladaki, Hammou Ait Blal, Tove Alm, Anna Asplund, Lars Björk, Lisa M Breckels, et al. A subcellular map of the human proteome. *Science*, 356(6340):eaal3321, 2017.
- Christopher P Toseland. Fluorescent labeling and modification of proteins. *Journal of chemical biology*, 6(3):85–95, 2013.
- Satoru Tsugawa, Nathan Hervieux, Oliver Hamant, Arezki Boudaoud, Richard S Smith, Chun-Biu Li, and Tamiki Komatsuzaki. Extracting subcellular fibrillar alignment with error estimation: application to microtubules. *Biophysical journal*, 110(8):1836–1844, 2016.
- Stefan Van der Walt, Johannes L Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D Warner, Neil Yager, Emmanuelle Gouillart, and Tony Yu. scikit-image: image processing in python. *PeerJ*, 2:e453, 2014.
- Hongda Wang, Yair Rivenson, Yiyin Jin, Zhensong Wei, Ronald Gao, Harun Günaydın, Laurent A Bentolila, Comert Kural, and Aydogan Ozcan. Deep learning enables cross-modality super-resolution in fluorescence microscopy. *Nat. Methods*, 16:103–110, 2019.
- Zhou Wang, Eero P Simoncelli, and Alan C Bovik. Multiscale structural similarity for image quality assessment. In *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, volume 2, pages 1398–1402. Ieee, 2003.
- Zhou Wang, Alan C Bovik, Hamid R Sheikh, Eero P Simoncelli, et al. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.
- Martin Weigert, Uwe Schmidt, Tobias Boothe, Andreas Müller, Alexandr Dibrov, Akanksha Jain, Benjamin Wilhelm, Deborah Schmidt, Coleman Broaddus, Siân Culley, et al. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nature methods*, 15(12):1090, 2018.
- Chih-Yuan Yang, Chao Ma, and Ming-Hsuan Yang. Single-image super-resolution: A benchmark. In *European Conference on Computer Vision*, pages 372–386. Springer, 2014.

Anne-Marie C Yvon, Patricia Wadsworth, and Mary Ann Jordan. Taxol suppresses dynamics of individual microtubules in living human tumor cells. *Molecular biology of the cell*, 10(4):947–959, 1999.

Hao Zhang, Chunyu Fang, Xinlin Xie, Yicong Yang, Wei Mei, Di Jin, and Peng Fei. High-throughput, high-resolution deep learning microscopy based on registration-free generative adversarial network. *Biomedical Optics Express*, 10(3):1044–1063, 2019.

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2223–2232, 2017.

Appendix

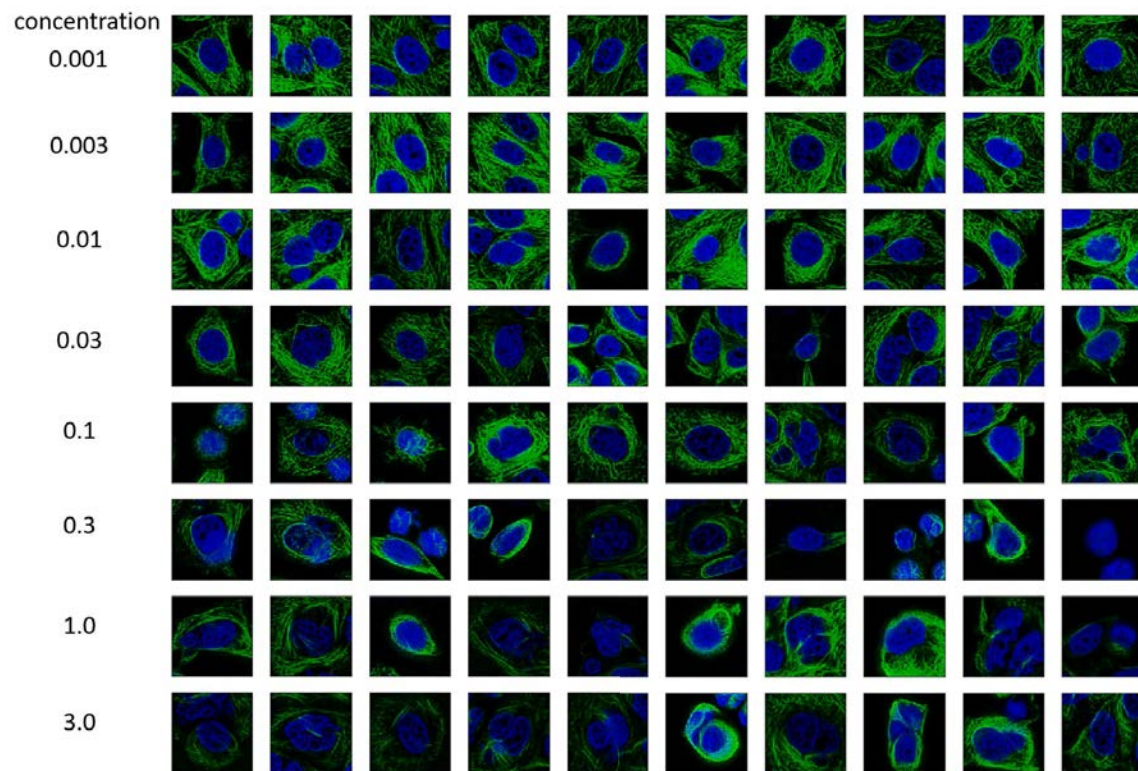


Figure 1: Examples of reconstructed images from cells treated with Etoposide B.

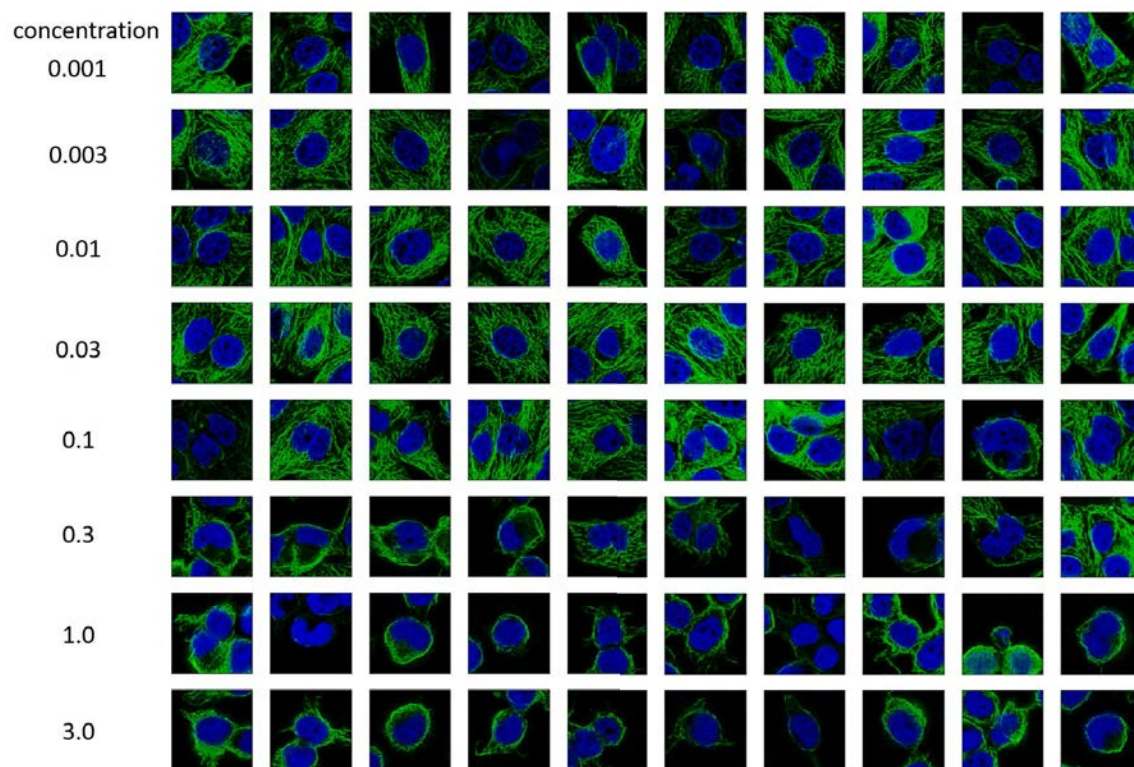


Figure 2: Examples of reconstructed images from cells treated with Latrunculin B.

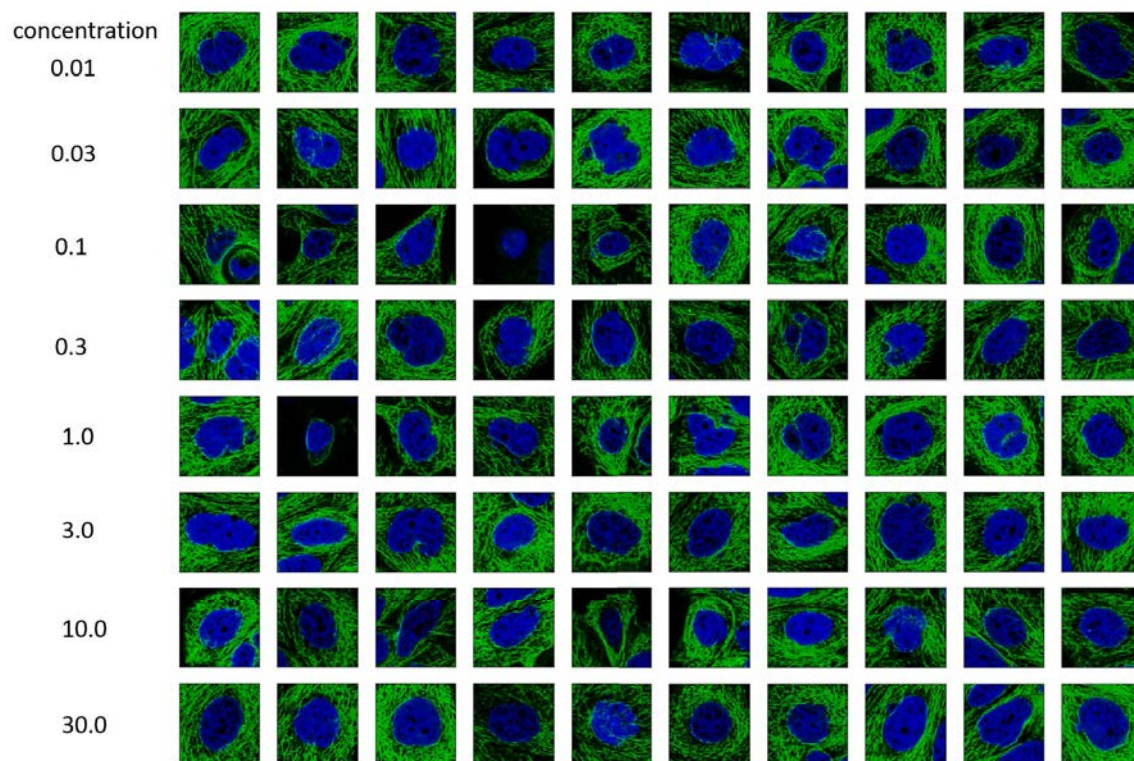


Figure 3: Examples of reconstructed images from cells treated with AZ-A.