

# DeepContrast: Deep Tissue Contrast Enhancement using Synthetic Data Degradations and OOD Model Predictions

Nuno Pimpão Martins<sup>1</sup>  
pimpaoma@mpi-cbg.de

Yannis Kalaidzidis<sup>1</sup>  
kalaidzi@mpi-cbg.de

Marino Zerial<sup>1</sup>  
zerial@mpi-cbg.de

Florian Jug<sup>1,2</sup>  
florian.jug@fht.org

<sup>1</sup> MPI-CBG, Dresden, Germany

<sup>2</sup> Fondazione Human Technopole, Milano, Italy

## Abstract

Microscopy images are crucial for life science research, allowing detailed inspection and characterization of cellular and tissue-level structures and functions. However, microscopy data are unavoidably affected by image degradations, such as noise, blur, or others. Many such degradations also contribute to a loss of image contrast, which becomes especially pronounced in deeper regions of thick samples. Today, best performing methods to increase the quality of images are based on Deep Learning approaches, which typically require ground truth (GT) data during training. Our inability to counteract blurring and contrast loss when imaging deep into samples prevents the acquisition of such clean GT data. The fact that the forward process of blurring and contrast loss deep into tissue can be modeled, allowed us to propose a new method that can circumvent the problem of unobtainable GT data. To this end, we first synthetically degraded the quality of microscopy images even further by using an approximate forward model for deep tissue image degradations. Then we trained a neural network that learned the inverse of this degradation function from our generated pairs of raw and degraded images. We demonstrated that networks trained in this way can be used out-of-distribution (OOD) to improve the quality of less severely degraded images, e.g. the raw data imaged in a microscope. Since the absolute level of degradation in such microscopy images can be stronger than the additional degradation introduced by our forward model, we also explored the effect of iterative predictions. Here, we observed that in each iteration the measured image contrast kept improving while detailed structures in the images got increasingly removed. Therefore, dependent on the desired downstream analysis, a balance between contrast improvement and retention of image details has to be found.

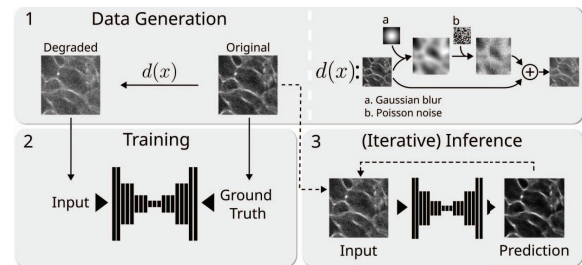


Figure 1. Proposed scheme to improve deep tissue contrast. (1) Pairs of data for supervised training are generated by degrading raw microscopy images using a suitable degradation function  $d(x)$  composed of a blurring and a noising step. (2) During supervised network training, synthetically degraded images are used as inputs and the original images as targets. (3) During inference, we feed the original raw microscopy images once or iteratively into the trained network (see Section 3).

## 1. Biological Motivation

In this work we applied our method (DEEPCONTRAST) to microscopy images of liver tissue. The liver is a frequently studied system in biomedical research, due to its vital functions in the human body, e.g. blood detoxification and bile production. Liver tissue is dense and compact, composed of many different cell types that display an intricate three-dimensional architecture. Still, many aspects of this structure are not fully understood, which drives biomedical research to use modern microscopy techniques to image large 3D sections of liver tissue at the highest achievable quality and resolution.

Data presented in this work was obtained using a Laser Scanning Confocal Microscope (LSM). This modality allowed us to obtain highly detailed image data in large and thick samples with sub-cellular resolution in all three spatial dimensions. Unfortunately, the image quality inevitably degrades in deeper layers of the imaged liver tissue, mostly due to light scattering. This poses a challenge to better our understanding of mesoscale structures that shape the liver in its full 3D complexity. Therefore, methods that facilitate

the downstream analysis of large 3D image data are much sought after.

## 2. Related Work

Classical algorithms to enhance contrast in images often rely on the intensity histogram, typically altering the overall histogram landscape with a set of predefined rules to either obtain a more uniformly sampled distribution or to match an histogram obtained from a desired reference. Examples to such algorithms are known as histogram matching [8] or histogram equalization [20]. These approaches are not content-aware, *i.e.*, they follow specific rules independent of the structures visible in the image to be modified. Contrast Limited Adaptive Histogram Equalization (CLAHE) [34] is an example of widely used histogram equalization method. We use this method as one of our baselines methods due to its popularity and widespread use in scientific image processing protocols.

Another popular family of algorithms used to improve image quality and contrast are feed-forward deconvolution methods such as the one by Richardson-Lucy [23, 14], or the popular Huygens software from [Scientific Volume Imaging](#). These are iterative approaches that attempt to undo the blurring induced by the point spread function (PSF) of the microscope. The main drawback of such approaches is the assumption of spatial invariance of the PSF, which does not hold in thick dense tissue microscopy.

Deep Learning (DL) based applications have proven to perform especially well on several image restoration tasks like denoising [29, 11, 3, 7, 22, 21], deconvolution [4, 9, 13], and super-resolution [19, 18, 26, 31, 32].

Content Aware Image Restoration (CARE) [29], uses supervised DL methods to restores microscopy image quality in various ways. However, in order to use CARE, it is necessary to obtain low and high quality versions of the same objects and structures, which is not possible in many real-world scenarios, such as the one presented in this work. One interesting insight with respect to out-of-distribution (OOD) denoising is presented in [15]. In it, the authors show that a network without trainable bias terms is more robust when applied to inputs that contain levels of noise that are inconsistent (OOD) with respect to the training data.

One popular way to solve the problem of GT data being required is to synthetically generate the required training pairs [28]. In [6], the authors used “crapified” images, as they call it, to obtain said training pairs for training super-resolution networks and networks that increase the temporal consistency in time-lapse movies. Others have used synthetic data generation for object detection [30] or segmentation [5].

Work specifically concerned with enhancing image contrast is less common. The FCE-Net [33] proposes a network architecture specifically designed to enhance image contrast

in biological image data. Since this makes FCE-Net our closest competitor, despite technically being a quite different approach, we will always also compare our own results to the ones obtainable with the FCE-Net.

## 3. Methods

Inspired by [29] and [6], we also set out to use the machinery of supervised learning in deep neural networks. In our case, for the sake of improving image contrast in microscopy data of large tissue samples. To this end, we synthetically generate appropriate training data, *i.e.* pairs of images that are of lower and higher contrast. Naturally, we cannot synthetically remove scattered light and noise from raw microscopy data, otherwise the very task we are seeking a solution for would be solved already. Instead, we can add additional light scattering and noise to the available raw data, making it even worse (see Figure 1).

More specifically, our degradation function is

$$d(x) = \alpha \cdot x + (1 - \alpha) \cdot n(b(x)), \quad (1)$$

where  $x$  is a raw input image,  $\alpha$  a hyperparameter that controls the blending between  $x$  and  $n(b(x))$ ,  $b$  is a blurring function that models light scattering in biological tissue, and  $n$  a function adding noise. In line with the most dominant noise in low-light fluorescent microscopy,  $n$  is adding Poisson noise to the blurred data.

Light scattering depends on refractive index transitions throughout the sample and a precise forward model is not easy to compute. For our purposes, the simple approximation introduced in Equation 1 leads to contrast enhancement results that outperform existing methods like the FCE-Net (see Section 4).

Once a body of input data  $X = (x_1, x_2, \dots, x_k)$  is further degraded to  $D = (d(x_1), \dots, d(x_k))$ , we use image pairs  $(d_i, x_i)$  for supervised training of a contrast enhancement network (Model A). The network was trained as a bias-free [15] U-Net [24]. The reason for *not* training the bias terms of the network nodes is that, once trained, it is intended to be applied to images  $x_i$  or similar, which are less severely affected by degradations, therefore, OOD with regards to  $d_i$ .

### Iterative Predictions

Since the absolute level of degradation in such microscopy images can be stronger than the additional degradation introduced by our forward model, we also explored the effect of iterative prediction. Iterative predictions with a trained DEEPCONTRAST network (DC) are simply multiple applications of DC to the input  $x$ . For example, the final DEEPCONTRAST (3×) prediction  $y$  is computed by

$$y = \text{DC}(\text{DC}(\text{DC}(x))). \quad (2)$$

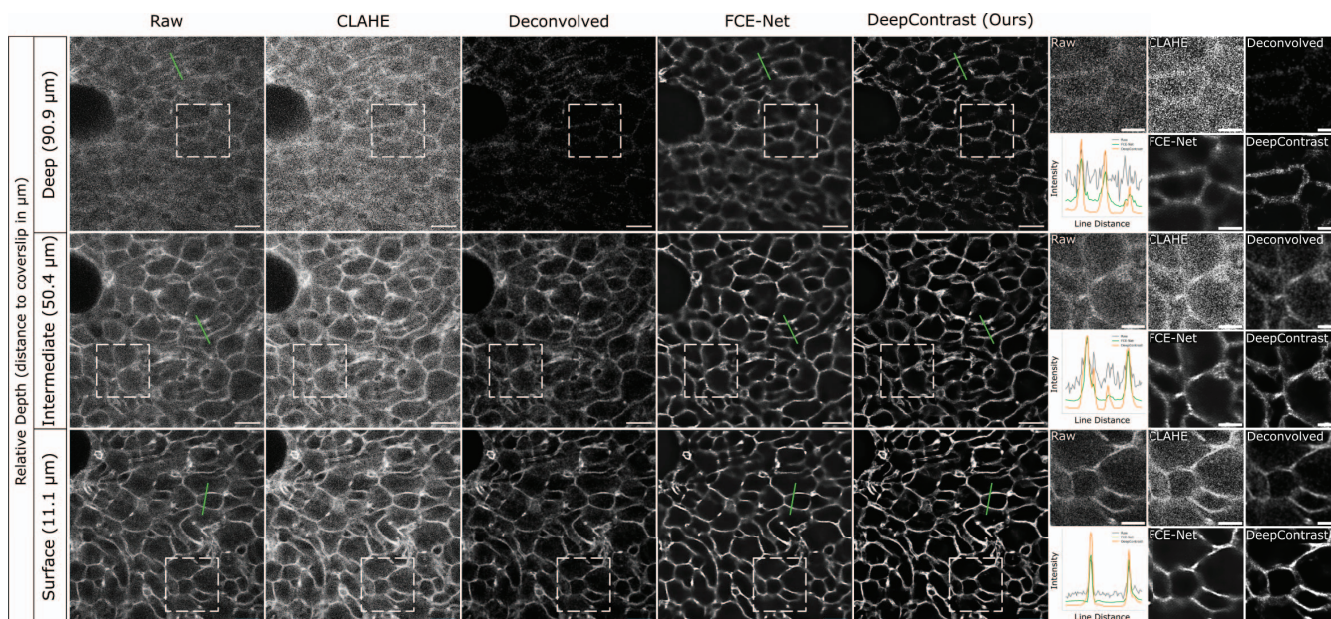


Figure 2. Qualitative results. Images of liver tissue sections stained with Phalloidin as a proxy for cell borders, used to compare our results (DEEPCONTRAST Model A) to several baseline methods. Rows show image planes at different depths in the liver tissue. Columns show the raw input, results obtained with CLAHE [34], Huygens deconvolution (see Section 4.4), best FCE-Net [33] results ( $3\times$ ), and our best DEEPCONTRAST results ( $3\times$ ), respectively. The three rightmost columns shows the inset areas marked by dashed boxes and line plots of raw intensities, the FCE-Net, and DEEPCONTRAST (along the green line in the respective images). Scale bars:  $20\mu\text{m}$  in full size images,  $10\mu\text{m}$  in insets.

The experiments we describe below and the results we show in Figure 3 indicate that iterative predictions indeed keep increasing image contrast. It must be noted, that better contrast does not necessarily mean that the predicted image is better for downstream analysis. Image faint details might get lost at the same time and the importance of such details depends on the downstream analysis to be conducted. We present one set of experiments on how to achieve a good balance between enhancing contrast and preserving image details in the following sections.

## 4. Experiments

### 4.1. Data

The imaged samples were cleared liver tissue sections, as described in [16], stained with Phalloidin 488 antibody to label the Actin cortex at all cell borders. Images were acquired in a Zeiss LSM 780 confocal microscope, with a Zeiss LCI Plan-Neofluar 63x 1.3 NA Gly/Water objective, a 488nm wavelength excitation laser, an emission window range of 489 – 551nm, and a pinhole size of  $1AU$ . Images were acquired with an isotropic voxel size of  $0.3\mu\text{m}$ . The maximum imaging depth was  $100\mu\text{m}$ .

### 4.2. Image Degradation Model

To compute synthetically degraded images, as described in Eq. 1, we first blur and noise each 2D slice (focal plane) using a Gaussian filter ( $\sigma = 20$  pixels) and Poisson noise at an estimated magnitude as described in [10] (using the image analysis software MotionTracking [17]). Synthetic images were then merged with the original images using  $\alpha$  values ranging in linear steps from 0.5 to 0.3, with 0.5 being used for the most superficial slice.

### 4.3. Network Architecture and Hyperparameters

Our network is a U-Net [24], using a depth of five, 32 initial feature channels, an MAE loss, and a linear function as the last activation layer. Our models were trained until convergence with an initial learning rate of  $4 \times 10^{-4}$  for a total of 450 epochs with 200 steps per epoch. A step uses a batch size of 16, of which each patch is a  $128 \times 128$  pixels crop from the body of training data. Networks were built using the CSBDeep toolbox [29] using Tensorflow 2.2.1. By default, and if not otherwise stated, we would not train the bias terms of each network node (-bias) to improve OOD predictions [15], as also described in Section 3.

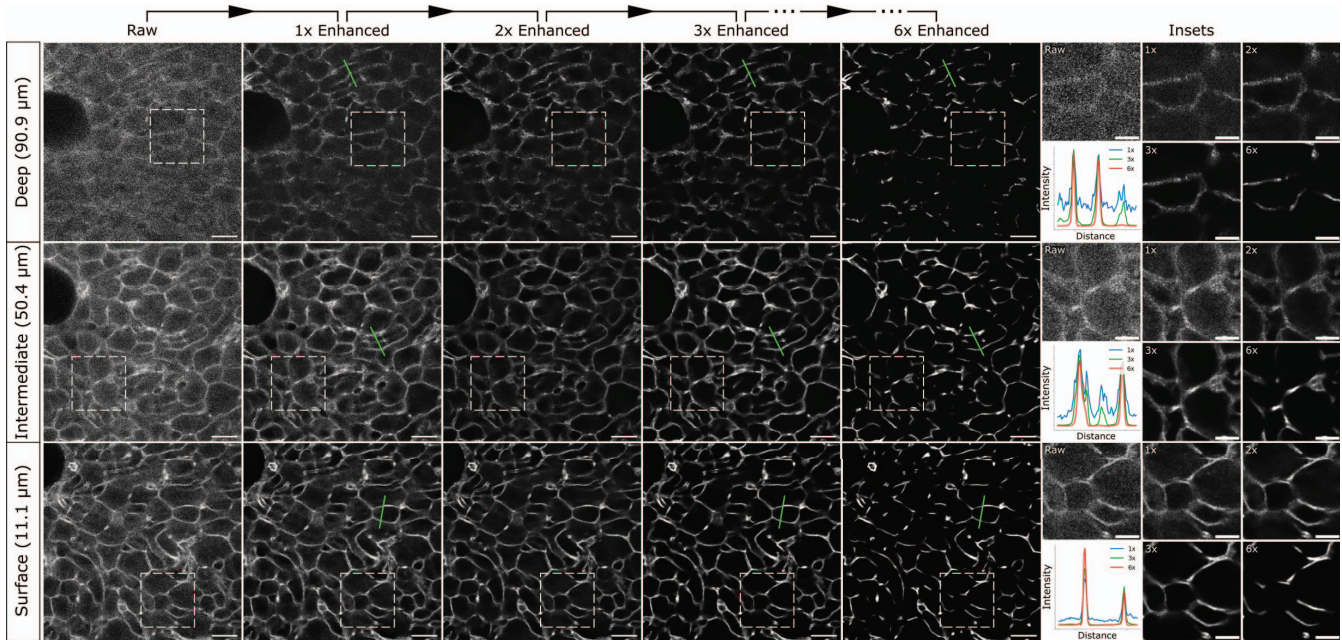


Figure 3. Qualitative results of iterative OOD model application. Contrast of the image data of Figure 2 iteratively enhanced using a trained DEEPCONTRAST network (Model A). Rows show, analogous to Figure 2, image planes at different depth into the imaged tissue. Columns show the raw input data, and the results of applying DEEPCONTRAST a single time, two, three (same as in Figure 2), and six consecutive times. The three rightmost columns shows the inset areas marked by dashed boxes and line plots along the green lines in the raw data, and along the 1 $\times$ , 3 $\times$ , and 6 $\times$  enhanced outputs. Note that, while contrast is continuously enhanced, too many iterative applications cause a notable loss of image details. Scale bars: 20  $\mu\text{m}$  in full size images, 10  $\mu\text{m}$  in insets.

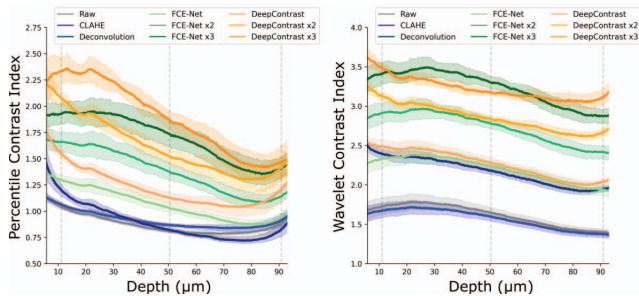


Figure 4. Quantitative results. Contrast quantification using the Percentile Contrast Index (see Section 4.7) and the Wavelet Contrast Index [1] (higher values are better) represented as average and 95% Confidence Intervals at each depth ( $N = 18$ ). Dashed vertical grey lines depict depths shown in Figure 2. Multiple iterations of FCE-Net [33] and our DEEPCONTRAST (Model A) approach show image contrast is further improved when these networks are iteratively applied.

#### 4.4. Baselines

Baseline methods, which we used to compare DEEPCONTRAST with, are: (i) classical methods, *i.e.* CLAHE and deconvolution using Huygens, and (ii) DL based methods, *i.e.* the FCE-Net [33] trained on our liver dataset.

CLAHE images were obtained using Fiji [25], where the *Enhanced Local Contrast* plugin is an implementation of the original CLAHE [34] method.

Deconvolved images were obtained using the software Huygens Professional (version 22.10.0p6) from Scientific Volume Imaging (SVI, The Netherlands), following provided pipelines with a theoretical PSF. Internally, Huygens is using the CMLE algorithm, with SNR set to 5, for 60 iterations, and the background value set to 100. This setup led to the best results on the data at hand.

For the FCE-Net [33] we used the code as it is provided by the authors of the original paper. Its worth mentioning that the provided pre-trained FCE network led to inferior results, hence, we trained the FCE-Net from scratch until convergence using our own data. Please note that we have also applied the FCE-Net iteratively, as described above in Equation 2 and observed that also FCE-Net results keep improving. For fair comparison we are therefore reporting iterative FCE-Net results whenever they are better than single predictions.

#### 4.5. Double Degradation Experiments

Since DEEPCONTRAST, by definition, is applied OOD, we wondered if iterative contrast enhancement will make consistent steps.

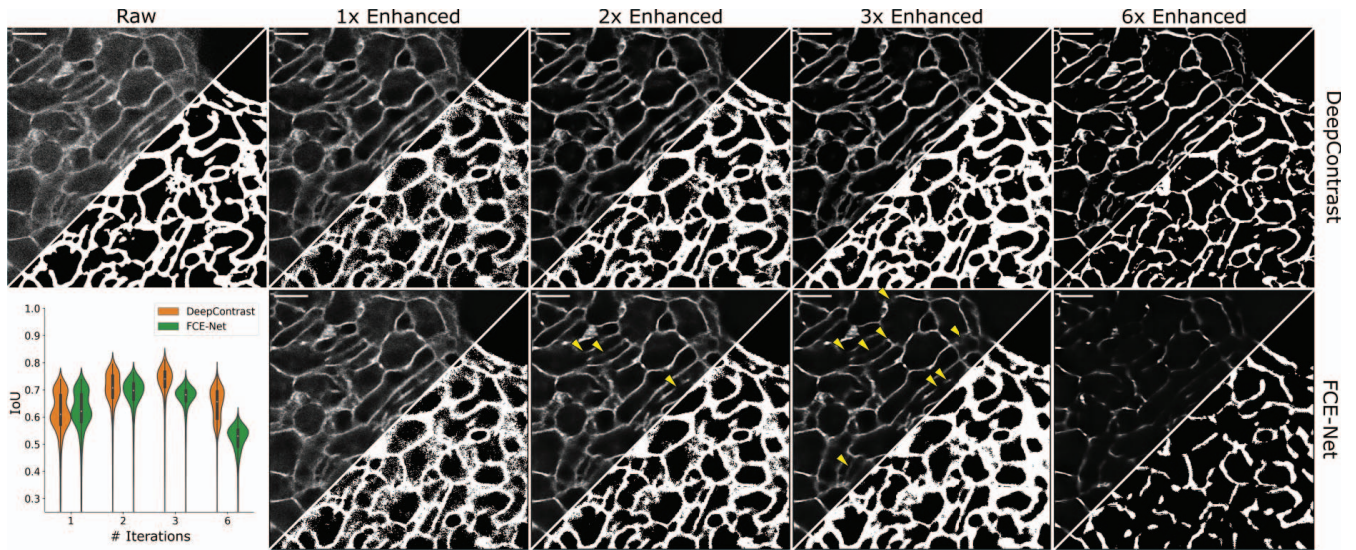


Figure 5. Qualitative and quantitative results of segmentation masks created at multiple iterations of contrast enhancement. Left side column shows Raw input and segmentation mask (Section 4.8). Each further column shows different iterations of contrast enhancement and corresponding segmentation of cell borders. Top row shows inference results with DEEPCONTRAST and bottom row shows inference results with FCE-Net. Yellow arrow heads in images highlight lost or degraded structures when comparing DEEPCONTRAST and FCE-Net. Violin plots shows distribution of IoU values between the different contrast enhancement methods and raw segmentation masks used as reference at multiple iterations ( $N = 2682$ ), showing faster decrease of IoU values and more abundant mistakes in segmentation with FCE-Net.

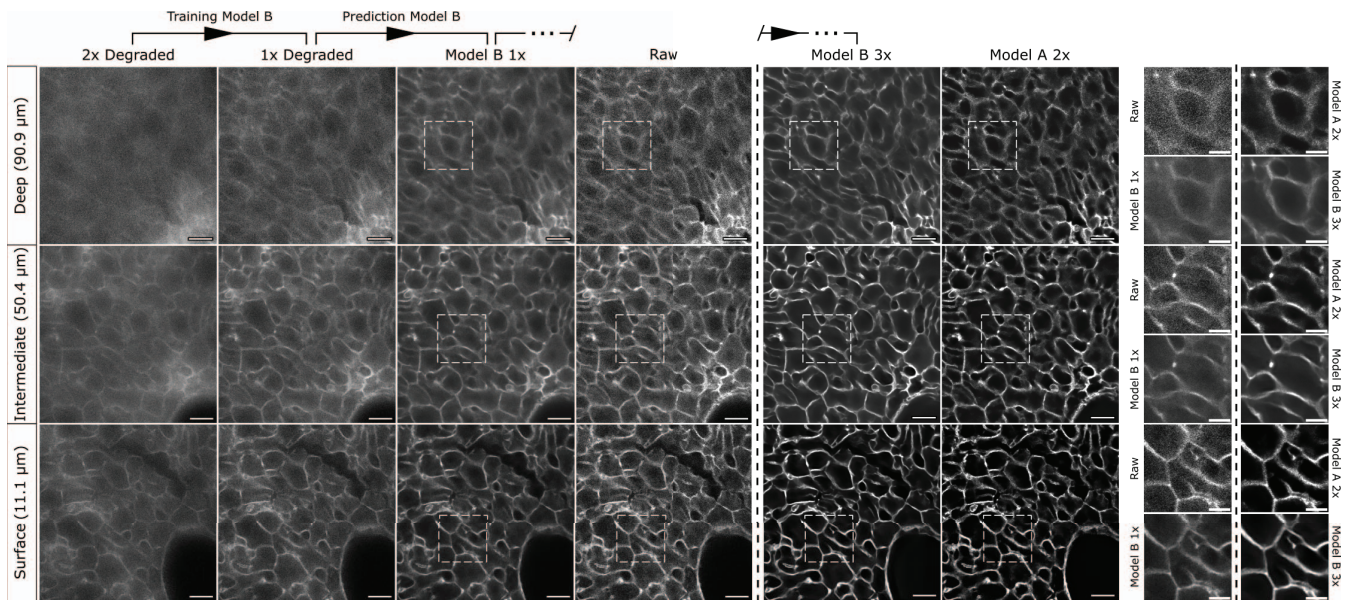


Figure 6. Qualitative double degradation results and predictions with Model B (see Section 4.5). Each row shows different depths in the sample, as in previous figures. Columns depict, from left to right, (i) double degraded images used as inputs during training, (ii) single degraded images used as targets during training, (iii) predictions of Model B when applied to data as the one in column two, (iv) the raw image data to compare Model B (1 $\times$ ) outputs against, (v) Model B (3 $\times$ ) outputs, and (vi) Model A (2 $\times$ ) outputs to compare Model B (3 $\times$ ) outputs against. Note that these two predictions should be and are similar since Model B starts with inputs that are once more synthetically degraded. The smaller panels in the rightmost columns show the insets (marked by dashed lines) from the other columns. Structures in all enhanced images are consistent with structures in the raw data, which is encouraging. Scale bars: 20  $\mu\text{m}$  in full size images, 10  $\mu\text{m}$  in insets.

	SSIM		
	$MB_{1\times}$ vs Raw	$MB_{2\times}$ vs $MA_{1\times}$	$MB_{3\times}$ vs $MA_{2\times}$
Very Deep	$0.51 \pm 0.07$	$0.70 \pm 0.08$	$0.80 \pm 0.07$
Deep	$0.52 \pm 0.07$	$0.72 \pm 0.09$	$0.82 \pm 0.07$
Intermediate	$0.56 \pm 0.06$	$0.77 \pm 0.06$	$0.84 \pm 0.03$
Shallow	$0.60 \pm 0.06$	$0.80 \pm 0.05$	$0.83 \pm 0.03$
Very Shallow	$0.59 \pm 0.08$	$0.79 \pm 0.06$	$0.83 \pm 0.02$

Table 1. Quantitative results of the double degradation experiment described in Section 4.5 and shown in Figure 6. We compare the outputs of three iterations of Model B ( $MB_{k\times}$ ), which was trained on double-degraded and single-degraded inputs, to the closest matching images, *i.e.* raw data for direct predictions of Model B ( $MB_{1\times}$  vs Raw), direct predictions of Model A to two iterations of Model B ( $MB_{2\times}$  vs  $MA_{1\times}$ ), and predictions of two iterations of Model A to three iterations of Model B ( $MB_{3\times}$  vs  $MA_{2\times}$ ). Section 3 for details.

As a first verification of our approach we degraded the raw image data twice, acquiring data triplets  $(e_i, d_i, x_i)$ , with  $x_i \in X$ , and  $d_i = d(x_i)$  and  $e_i = d(d(x_i))$ . Then we trained a DEEPCONTRAST network, Model B, on pairs  $(e_i, d_i)$  and applied the trained network, in-line with the initially proposed procedure, to  $d_i$  to increase its contrast and yielding  $y_i = DC(d_i)$ .

Since we started by double degrading the original  $x_i$ , we can now compare the prediction  $y_i$  with  $x_i$ , and further iterations of Model B with corresponding predictions obtained with Model A (see Section 3). If the trained DEEPCONTRAST network is indeed a good approximation of the inverse of our degradation function  $d$ , predictions  $y_i$  should be similar to the original images  $x_i$  at the first iteration, and to the corresponding iteration of output images from Model A.

#### 4.6. Ablations

In order to evaluate if bias-free training [15] is indeed leading to better results, we decided to repeat model training also on networks that are not bias-free, *i.e.* train all weights and biases.

#### 4.7. Contrast Quantification

##### Wavelet Contrast Index

With increasing contrast in an image, we expect background signal to be reduced and, consequently, the brightness of biological structures in the image (signal) to be increased. To quantify image contrast when no GT data is available, we used the Wavelet Contrast Index (WCI) [1]. This measure computes the difference between coefficients obtained from a wavelet decomposition, following the equation

$$WCI(x) = \log\left(\frac{W_{95^{th}}(x)}{W_{50^{th}}(x)}\right), \quad (3)$$

where  $x$  is the input image for which we want to evaluate the contrast,  $W_{95^{th}}$  is the 95<sup>th</sup> percentile wavelet coefficient and  $W_{50^{th}}$  is the median wavelet coefficient.

Wavelet decomposition was performed with the Py-Wavelets [12] python package using a Haar wavelet as the reference function and used coefficients up to the fourth level of decomposition.

##### Percentile Contrast Index

We also used the Percentile Contrast Index (PCI) to quantify intensity differences between image background and image structures. The PCI is computed by

$$PCI(x) = \log\left(\frac{I_{95^{th}}(x)}{I_{50^{th}}(x)}\right), \quad (4)$$

where  $x$  is again the input image to evaluate, and  $I_{95^{th}}$  is the 95<sup>th</sup> intensity value in the image being analyzed and  $I_{50^{th}}$  is the median intensity of  $x$ . We use the median value, assuming that at least half the pixels of any given image are background pixels.

#### 4.8. Downstream Segmentation after Contrast Enhancement

Enhancing contrast not necessarily improves downstream process-ability (interpretability) of a given dataset. While the contrast, as measured by WCI and/or PCI, might still improve, details relevant for biological interpretation of the data might already get lost. Therefore, the best amount of contrast enhancement depends on a given downstream analysis task. To this end, we introduced a downstream segmentation task and checked if a fixed segmentation pipeline improved with respect to existing ground truth labels. GT segmentation masks were generated from raw data using Labkit [2] (available as a Fiji [25] plugin).

For simplicity, we segmented contrast enhanced images  $y_i$  by thresholding, optimizing for the best threshold value, *i.e.* the one that maximizes the intersection-over-union (IoU) with respect to the previously generated GT. Our reasoning was that enhancing contrast would result in a better IoU after thresholding as long as relevant structures in the contrast enhanced images  $y_i$  were not lost. As soon as details were getting lost, the IoU dropped, allowing us to choose the most sensible iteration depth for DEEPCONTRAST (or the FCE-Net).

### 5. Results

Qualitative results presented in Figure 2 suggest that DEEPCONTRAST outperforms all baseline methods. DEEPCONTRAST removes or reduces image noise and enhances the intensity of visible image structures seemingly without losing fine details (signal) from predicted images.

Hence, DEEPCONTRAST is indeed increasing image contrast.

Both classical methods, *i.e.* CLAHE and deconvolution, displayed relatively poor results mainly deep into the tissue. CLAHE amplified image noise at all imaging depths and mostly failed to highlight biological structures. Deconvolution, on the other hand, did reduce image noise, but failed to increase intensities of foreground structures (most obvious deep into the tissue).

The FCE-Net performed much better, leading to good results close to the surface. But the image contrast in FCE-Net predictions decayed with increasing depth (see insets in Figure 2). Qualitatively, the FCE-Net also seemed to produce less sharp cell borders (as seen in either deep and shallow image regions).

To validate these qualitative observations, we quantified contrast with the two measures WCI and PCI (see Section 4.7). As can be seen in Figure 4, DEEPCONTRAST achieved higher image contrast at all imaging depths and over all plotted iterative applications ( $1\times$  to  $3\times$ ). One notable exception are the WCI values for  $3\times$  iterations in intermediate imaging depths. In these images, despite the FCE-Net showing higher WCI values, one can see more image structure being lost in FCE-Net predictions than in predictions obtained with DEEPCONTRAST (see Figure 5 for a qualitative and quantitative comparison).

As introduced above, contrast enhancement can be applied iteratively (Equation 2). Results of performing multiple rounds of enhancement are shown in Figure 3. Visually, the best results were obtained with three rounds of enhancement ( $3\times$ ). While contrast readouts using WCI and PCI would still improve with additional iterations, image details would start disappearing (as can be seen in the  $6\times$  column and the line-plots in Figure 3).

Qualitative results of the Double Degradation Experiments introduced in Section 4.5, are shown in Figure 6. Predictions of Model B at iteration  $k$  should and are corresponding well to predictions of Model A at iteration  $k - 1$  since Model B is trained on image pairs that are one application of our forward degradation model (d) more degraded. We quantify this via structure similarity index measure (SSIM) [27] in Table 1 and allow for a qualitative comparison between the corresponding columns in Figure 6.

### 5.1. Contrast Enhancement vs. Segmentation

To better quantify the undesired effect of losing relevant details while simultaneously gaining additional contrast in the processed microscopy data, we introduced a simple threshold based segmentation task (See Section 4.8). A qualitative as well as quantitative comparison is shown in Figure 5. IoU values are initially increasing with number of iterations, but then eventually drop when too many image structures are removed. The FCE-Net generally shows

lower IoU values, suggesting that DEEPCONTRAST is not only leading to more contrasted images, but is at the same time maintaining more image details with iterations. In addition to the IoU quantification, we highlighted lost details on images with yellow arrow heads (see in Figure 5), pointing differences between iterative inferences.

### 5.2. Ablation: Training including Bias

As introduced above, DEEPCONTRAST employs bias-free [15] network training. In Figure 7 we show representative predictions of Model A ( $3\times$ ), as used in Figure 2, and compare them to predictions obtained with an equivalent model which was trained with bias (+bias). Yellow arrow heads in the figure point at locations where the bias free network does a better job retaining image details. Empirically, we did not spot any cases where the opposite is true, which gave us additional motivation to use bias-free networks in DEEPCONTRAST.

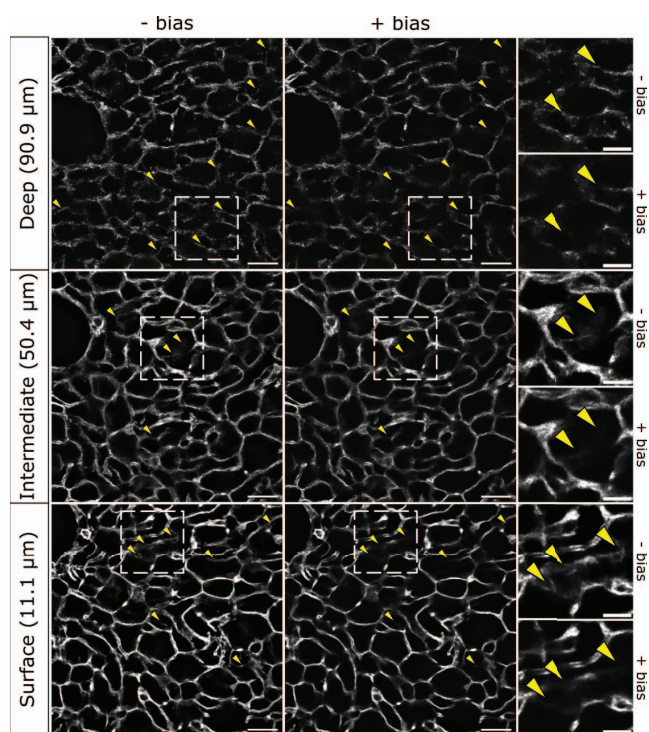


Figure 7. Qualitative results of networks trained without and with bias. Phalloidin stained images of liver tissue sections enhanced  $3\times$  with DEEPCONTRAST models trained with (+bias; right side) and without bias (-bias; left side). Rows show image planes at different depths relative to cover-glass. Network model trained with bias performs worse when applied OOD, removing structures seen in a model trained without bias (- bias), highlighted by yellow arrow-heads. Scale bars:  $20\ \mu\text{m}$  in full size images,  $10\ \mu\text{m}$  in insets.

## 6. Discussion and Conclusion

In this work we propose to use an image degradation function to approximate light scattering in deep tissue imaging and use it to generate synthetically degraded data to enable supervised network training. Our results show that the relatively simple degradation model we introduced is sufficient to increase image contrast in real microscopy data. Our method can be applied in an iterative manner to further increase image contrast and will retain detailed image structures for more iterations than the competitive baseline methods we compared against.

For the liver data at hand, we have found that the best number of iterations for contrast enhancement is three ( $3\times$ ). This assessment is based on a combination of contrast enhancement and retention of fine image details in the contrast enhanced predictions. A more quantitative approach to the visual assessment was introduced by means of a downstream segmentation task, which has indeed confirmed our initial findings.

In general, the best trade-off between contrast enhancement and structural integrity of predictions depends on the nature of the downstream processing task to be conducted. Hence, an analysis similar to the one we performed for the segmentation task could be required to evaluate the best-performing setup.

Similarly we found that for better OOD application of our trained networks, the bias free version seems to lead to better results.

While our approach is leading to excellent results and can easily be used by microscopists and life scientists to improve volumetric image data for quantitative downstream processing, additional research will be required to undo image degradations deep in imaged tissues in more fundamental ways.

## Acknowledgments

The authors thank José Valenzuela-Iturra for acquiring image data. We thank the LMF and SCF facilities at MPI-CBG for technical support. We thank Igor Zubarev for help and feedback with the presented experimental setups. We thank Ashesh, Anirban Ray, Sheida Rahnamai Kordasiabi, Igor Zubarev, Joran Deschamps and Damian Dalle Nogare for helpful discussion and feedback. This work was supported by the European Research Council ERC Advanced Rulliver Grant (no. 695646) to M. Zerial. Additionally, this work was supported by the European Commission through the Horizon Europe program AI4LIFE with grant agreement 101057970-AI4LIFE. Funding was also provided from the Max-Planck Society under project code M.IF.A.MOZG8106.

## References

- [1] Anna Lea Albright and Peter Huybers. Paintings by Turner and Monet depict trends in 19th century air pollution. *Proceedings of the National Academy of Sciences*, 120(6):e2219118120, Feb. 2023. Publisher: Proceedings of the National Academy of Sciences. 4, 6
- [2] Matthias Arzt, Joran Deschamps, Christopher Schmied, Tobias Pietzsch, Deborah Schmidt, Pavel Tomancak, Robert Haase, and Florian Jug. LABKIT: Labeling and Segmentation Toolkit for Big Image Data. *Frontiers in Computer Science*, 4, 2022. 6
- [3] Joshua Batson and Loic Royer. Noise2Self: Blind Denoising by Self-Supervision. *36th International Conference on Machine Learning, ICML 2019*, 2019-June:826–835, Jan. 2019. Publisher: International Machine Learning Society (IMLS) .eprint: 1901.11365. 2
- [4] Jiji Chen, Hideki Sasaki, Hoyin Lai, Yijun Su, Jiamin Liu, Yicong Wu, Alexander Zhovmer, Christian A. Combs, Ivan Rey-Suarez, Hung-Yu Chang, Chi Chou Huang, Xuesong Li, Min Guo, Srineil Nizambad, Arpita Upadhyaya, Shih-Jong J. Lee, Luciano A. G. Lucas, and Hari Shroff. Three-dimensional residual channel attention networks denoise and sharpen fluorescence microscopy image volumes. *Nature Methods*, 18(6):678–687, June 2021. Number: 6 Publisher: Nature Publishing Group. 2
- [5] Kenneth W. Dunn, Chichen Fu, David Joon Ho, Soonam Lee, Shuo Han, Paul Salama, and Edward J. Delp. Deep-Synth: Three-dimensional nuclear segmentation of biological images using neural networks trained with synthetic data. *Scientific Reports*, 9(1):18295, Dec. 2019. Number: 1 Publisher: Nature Publishing Group. 2
- [6] Linjing Fang, Fred Monroe, Sammy Weiser Novak, Lyndsey Kirk, Cara R. Schiavon, Seungyoon B. Yu, Tong Zhang, Melissa Wu, Kyle Kastner, Alaa Abdel Latif, Zijun Lin, Andrew Shaw, Yoshiyuki Kubota, John Mendenhall, Zhao Zhang, Gulcin Pekkurnaz, Kristen Harris, Jeremy Howard, and Uri Manor. Deep learning-based point-scanning super-resolution imaging. *Nature Methods*, 18(4):406–416, Apr. 2021. Number: 4 Publisher: Nature Publishing Group. 2
- [7] Anna S. Goncharova, Alf Honigmann, Florian Jug, and Alexander Krull. Improving Blind Spot Denoising for Microscopy, Aug. 2020. arXiv:2008.08414 [cs, eess]. 2
- [8] RC Gonzalez and RE Woods. Digital image processing third edition, str. 653–658, 2006. 2
- [9] Min Guo, Yue Li, Yijun Su, Talley Lambert, Damian Dalle Nogare, Mark W. Moyle, Leighton H. Duncan, Richard Ikegami, Anthony Santella, Ivan Rey-Suarez, Daniel Green, Anastasia Beiriger, Jiji Chen, Harshad Vishwasrao, Sundar Ganesan, Victoria Prince, Jennifer C. Waters, Christina M. Annunziata, Markus Hafner, William A. Mohler, Ajay B. Chitnis, Arpita Upadhyaya, Ted B. Usdin, Zhirong Bao, Daniel Colón-Ramos, Patrick La Riviere, Huafeng Liu, Yicong Wu, and Hari Shroff. Rapid image deconvolution and multiview fusion for optical microscopy. *Nature Biotechnology*, 38(11):1337–1346, Nov. 2020. Number: 11 Publisher: Nature Publishing Group. 2



- [10] Yannis Kalaidzidis. Fluorescence Microscopy Noise Model : Estimation of Poisson Noise Parameters from Snap-Shot Image. *BIOCOMP'17*, (1):63–66, 2017. ISBN: 1601324502. 3
- [11] Alexander Krull, Tim Oliver Buchholz, and Florian Jug. Noise2void-Learning denoising from single noisy images. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2019-June:2124–2132, 2019. ISBN: 9781728132938 \_eprint: 1811.10980. 2
- [12] Gregory R. Lee, Ralf Gommers, Filip Waselewski, Kai Wohlfahrt, and Aaron O’Leary. PyWavelets: A Python package for wavelet analysis. *Journal of Open Source Software*, 4(36):1237, Apr. 2019. 6
- [13] Yue Li, Yijun Su, Min Guo, Xiaofei Han, Jiamin Liu, Harshad D. Vishwasrao, Xuesong Li, Ryan Christensen, Titas Sengupta, Mark W. Moyle, Ivan Rey-Suarez, Jiji Chen, Arpita Upadhyaya, Ted B. Usdin, Daniel Alfonso Colón-Ramos, Huafeng Liu, Yicong Wu, and Hari Shroff. Incorporating the image formation process into deep learning improves network performance. *Nature Methods*, 19(11):1427–1437, Nov. 2022. Number: 11 Publisher: Nature Publishing Group. 2
- [14] L. B. Lucy. An iterative technique for the rectification of observed distributions. *The Astronomical Journal*, 79:745, June 1974. ADS Bibcode: 1974AJ.....79..745L. 2
- [15] Sreyas Mohan, Zahra Kadkhodaie, Eero P. Simoncelli, and Carlos Fernandez-Granda. Robust and interpretable blind image denoising via bias-free convolutional neural networks, Feb. 2020. arXiv:1906.05478 [cs, eess, stat]. 2, 3, 6, 7
- [16] Hernán Morales-Navarrete, Hidenori Nonaka, André Scholich, Fabián Segovia-Miranda, Walter de Back, Kirstin Meyer, Roman L Bogorad, Victor Koteliensky, Lutz Bruschi, Yannis Kalaidzidis, Frank Jülicher, Benjamin M Friedrich, and Marino Zerial. Liquid-crystal organization of liver tissue. *eLife*, 8:1–18, 2019. 3
- [17] Hernán Morales-Navarrete, Fabián Segovia-Miranda, Piotr Klukowski, Kirstin Meyer, Hidenori Nonaka, Giovanni Marsico, Mikhail Chernykh, Alexander Kalaidzidis, Marino Zerial, and Yannis Kalaidzidis. A versatile pipeline for the multi-scale digital reconstruction and quantitative analysis of 3D tissue architecture. *eLife*, 4(DECEMBER2015):1–29, 2015. 3
- [18] Elias Nehme, Lucien E Weiss, Tomer Michaeli, and Yoav Shechtman. Deep-STORM: super-resolution single-molecule microscopy by deep learning. *Optica*, 5(4):458, Apr. 2018. 2
- [19] Wei Ouyang, Andrey Aristov, Mickaël Lelek, Xian Hao, and Christophe Zimmer. Deep learning massively accelerates super-resolution localization microscopy. *Nat. Biotechnol.*, 27:2672, Apr. 2018. 2
- [20] Stephen M. Pizer, E. Philip Amburn, John D. Austin, Robert Cromartie, Ari Geselowitz, Trey Greer, Bart ter Haar Romeny, John B. Zimmerman, and Karel Zuiderveld. Adaptive histogram equalization and its variations. *Computer Vision, Graphics, and Image Processing*, 39(3):355–368, Sept. 1987. 2
- [21] Mangal Prakash, Mauricio Delbracio, Peyman Milanfar, and Florian Jug. Interpretable Unsupervised Diversity Denoising and Artefact Removal, Feb. 2022. arXiv:2104.01374 [cs, eess, q-bio]. 2
- [22] Mangal Prakash, Alexander Krull, and Florian Jug. Fully Unsupervised Diversity Denoising with Convolutional Variational Autoencoders, Mar. 2021. arXiv:2006.06072 [cs, eess]. 2
- [23] William Hadley Richardson. Bayesian-Based Iterative Method of Image Restoration\*. *JOSA*, 62(1):55–59, Jan. 1972. Publisher: Optica Publishing Group. 2
- [24] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 9351, pages 234–241. Springer Verlag, 2015. ISSN: 16113349 \_eprint: 1505.04597. 2, 3
- [25] Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, Jean-Yves Tinevez, Daniel James White, Volker Hartenstein, Kevin Eliceiri, Pavel Tomancak, and Albert Cardona. Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7):676–682, July 2012. Number: 7 Publisher: Nature Publishing Group. 4, 6
- [26] Ying Tai, Jian Yang, and Xiaoming Liu. Image Super-Resolution via Deep Recursive Residual Network. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2790–2798, Honolulu, HI, July 2017. IEEE. 2
- [27] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image Quality Assessment: From Error Visibility to Structural Similarity. Technical Report 4, 2004. Publication Title: IEEE TRANSACTIONS ON IMAGE PROCESSING Volume: 13. 7
- [28] Martin Weigert, Loic Royer, Florian Jug, and Gene Myers. Isotropic reconstruction of 3D fluorescence microscopy images using convolutional neural networks, Apr. 2017. arXiv:1704.01510 [cs]. 2
- [29] Martin Weigert, Uwe Schmidt, Tobias Boothe, Andreas Müller, Alexandr Dibrov, Akanksha Jain, Benjamin Wilhelm, Deborah Schmidt, Coleman Broaddus, Siân Culley, Mauricio Rocha-Martins, Fabián Segovia-Miranda, Caren Norden, Ricardo Henriques, Marino Zerial, Michele Solimena, Jochen Rink, Pavel Tomancak, Loic Royer, Florian Jug, and Eugene W. Myers. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nature Methods*, 15(12):1090–1097, Dec. 2018. Publisher: Nature Publishing Group \_eprint: 236463. 2, 3
- [30] Ruijie Yao, Jiaqiang Qian, and Qiang Huang. Deep-learning with synthetic data enables automated picking of cryo-EM particle images of biological macromolecules. *Bioinformatics*, 36(4):1252–1259, Feb. 2020. 2
- [31] Yulun Zhang, Kungpeng Li, Kai Li, Lichen Wang, Bineng Zhong, and Yun Fu. Image Super-Resolution Using Very Deep Residual Channel Attention Networks, July 2018. arXiv:1807.02758 [cs] version: 2. 2
- [32] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu. Residual Dense Network for Image Super-

Resolution, Mar. 2018. arXiv:1802.08797 [cs] version: 2.  
2

- [33] Yunfei Zhang, Peng Wu, Siqi Chen, Hui Gong, and Xiaoquan Yang. FCE-Net: a fast image contrast enhancement method based on deep learning for biomedical optical images. *Biomedical optics express*, 13(6):3521–3534, June 2022. 2, 3, 4
- [34] K Zuiderveld. Contrast Limited Adaptive Histogram Equalization. *Graphics Gems*, 0:474–485, 1994. Publisher: Academic Press. 2, 3, 4