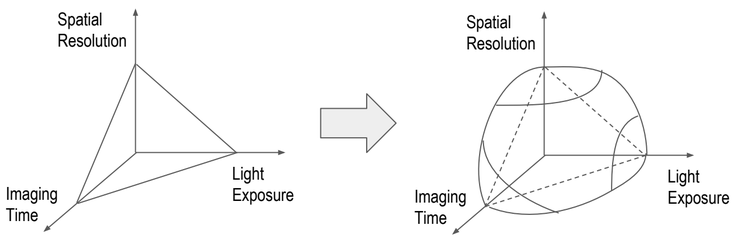
Fluorescent Microscopy Image Restoration

Colin Andrus, Dustin Godevais, Hollis Nymark, Joe Sloan

Introduction

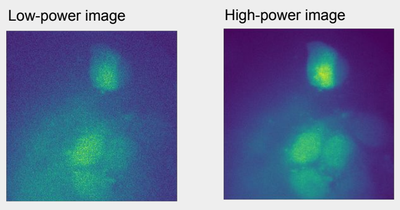
Problem statement: Advanced fluorescence microscopes use high-powered LEDs and lasers to observe cells over time. However, observational cells will eventually degrade under extended light exposure resulting in unusable, damaged cells.

Goal: Expand the tradeoff boundary between spatial resolution and exposure time using techniques from Deep Learning.



- Similar to super-resolution and image recovery tasks.
- Expand the possibilities frontier.

Motivation: Study RNA distribution and gene expression at lower power channels.



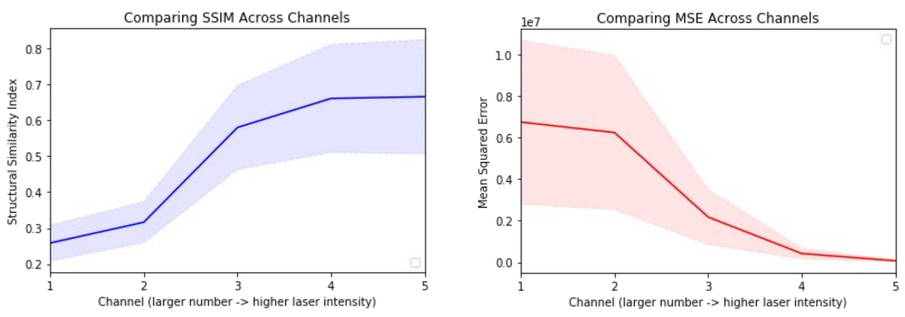
Data Overview

Data Overview: Stacks of 512 x 512 images taken by a fluorescence microscope across 7 different channels at each focal distance.

Data collected by an automated process using a fluorescence microscope on cells stained with DAPI and GFP:

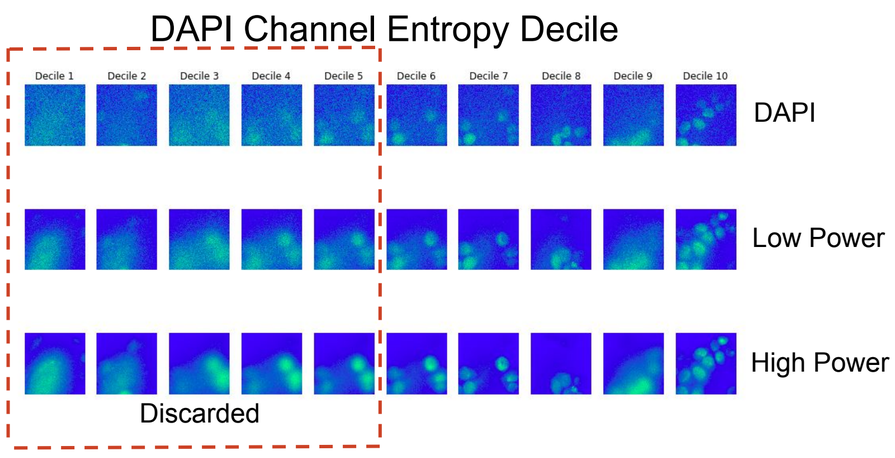
- DAPI highlights cell nuclei
- GFP tracks mRNA distribution

Each sample consists of a stack of 100 focal planes (z-slices) with 7 channels at each slice (1 for DAPI, 6 at varying power for GFP).

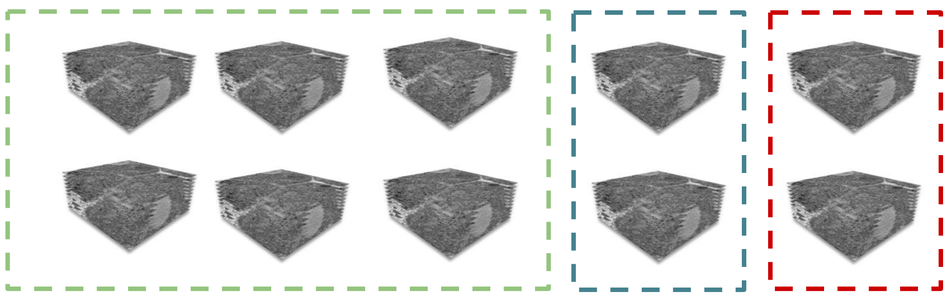


Data Selection: Due to microscope focus issues, many z-slices contain fuzzy images that are of low quality in all channels, regardless of laser intensity.

We select image sets with suitable clarity using the entropy of pixel intensities for an additional DAPI stained channel.



Data Preprocessing: After selection, we split data into train/validation/test by stack because we expect adjacent z-planes to be similar enough to cause leakage issues if separated. In addition, data was augmented and normalized using standard techniques.



Representative diagram of splitting, with green, blue, and red as train, validation, and test respectively. Splitting by z-slice would lead to significant leakage.

Evaluation Methods

MSE was generally used as the training loss (unless certain baselines had other loss functions for specific components) while evaluation was done with Structural Similarity Index (SSIM):

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

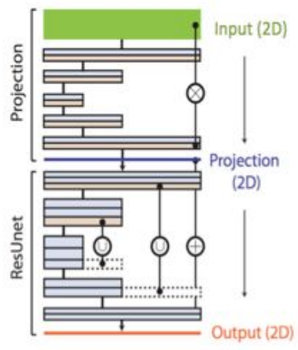
Here, μ and σ are the mean and variance (or covariance) respectively of the distribution of pixel intensity in each image. The constants are for numerical stability and generally a small value scaled with intensity range. This is computed for a small patch at a time.

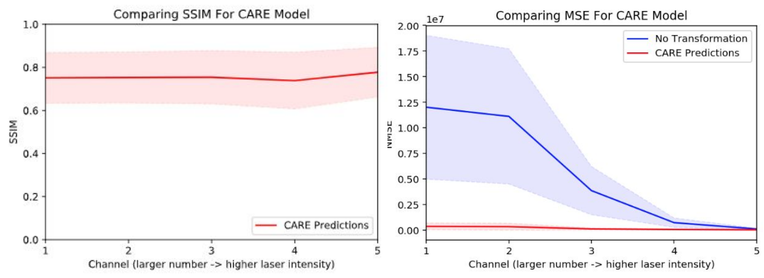
- Note that evaluation methods such as SSIM and MSE are *not scale invariant* to the image data type (e.g. int16, uint32, float, etc.). Thus, we convert our data image type to float for the SSIM comparison.
- The reported SSIM values are an average of this score over corresponding patches of the two images. SSIM lies in the range [-1,1] with a score of 1 implying identical images.

Data Type	Data Range
int8	0 to 255
int16	-32768 to 32767
float	0 to 1
uint32	0 to 2^31 - 1

Baselines

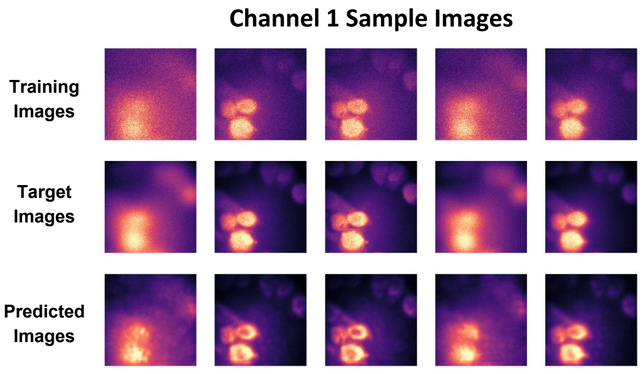
CSBDeep Model





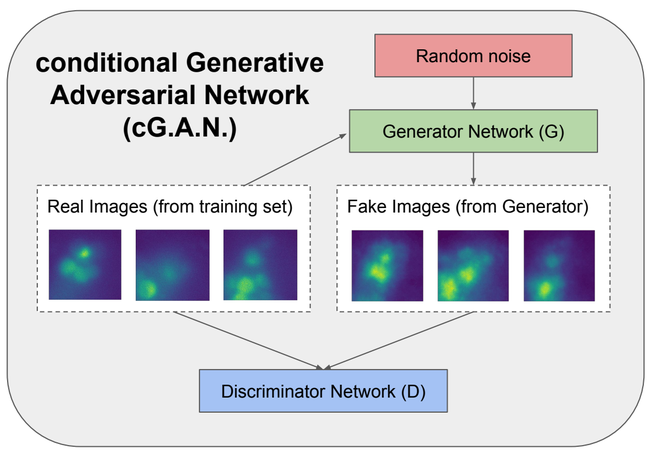
The CARE network (content-aware image restoration network) consists of two consecutive sub-networks: the first on projects voxel intensities, while the second denoises the projection. Outputs represent the mean prediction of a Laplace distribution for each pixel.

$$g_\theta(x)_i = p_{\text{laplace}}\big(\mu_\theta(x)_i, \sigma_\theta(x)_i\big)$$

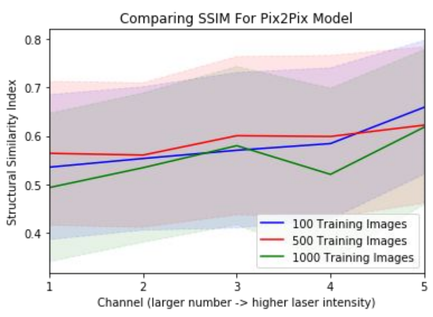


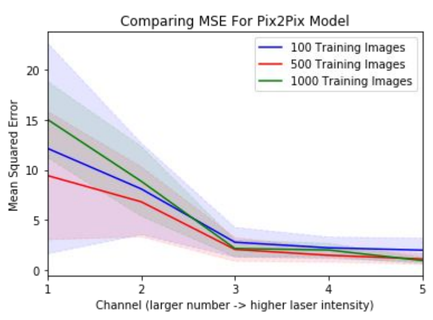
*Note on Image Color: CARE util functions recolor images (see normal output in conclusion)

Pix2Pix Model



$$G^* = \arg\min_G \max_D \mathcal{L}_{cGAN}(G,D) + \lambda \mathcal{L}_{L1}(G).$$

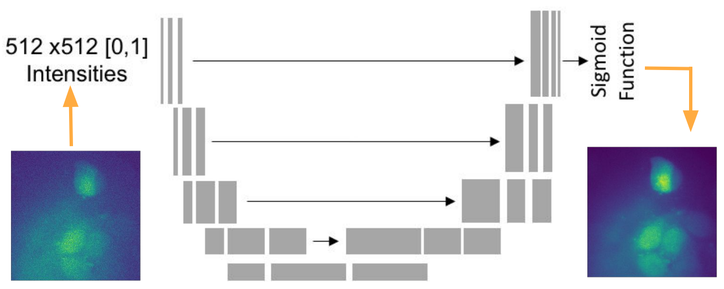




Learning Rate Policy*	Validation SSIM	Validation MSE
Step (gamma=0.1)	0.57069	10.2457
Lambda	0.56401	9.42761
Cosine Annealing	0.57204	10.7071

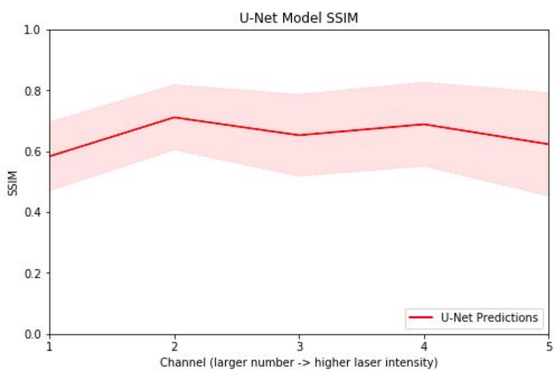
* This hyperparameter tuning of cGAN was performed with training sample = 500

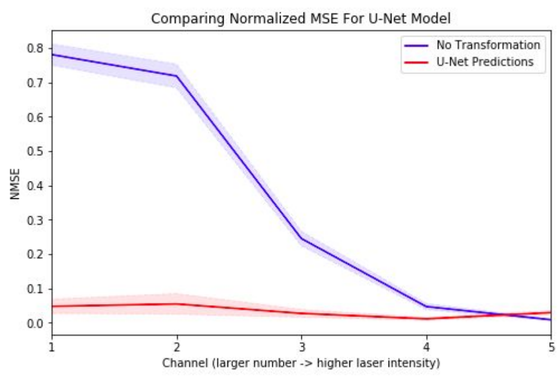
U-Net Model



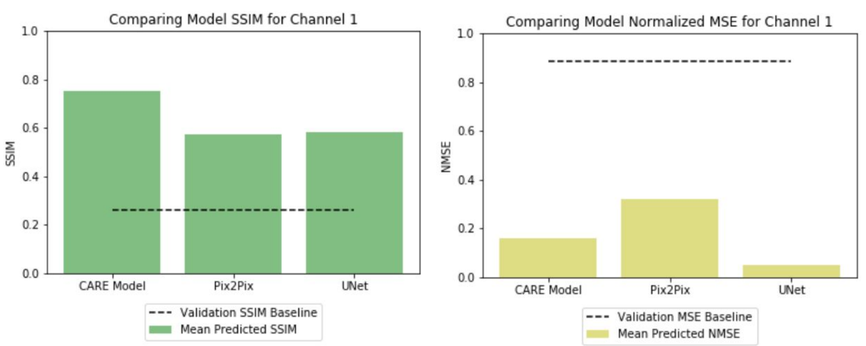
The U-Net architecture uses a series of down convolution, ReLU and max pooling steps which contract dimensions followed by series of deconvolutions, concatenations and convolutions that expand back to the original 512x512 dimension of the input.

The concatenations connecting the down and up paths allow local information to be used with the contextual information from the higher channel, lower dimensional steps.





Conclusion



For future applications, predictions based on Channel 1 (lowest power) provides the most value to end users.

Based on the mean validation set SSIM for Channel 1, the CARE Model performs best with a mean SSIM of 0.75.