Fine-tuning of Gaussian Denoising CNNs for Image Reconstruction in Fluorescence Microscopy

Leopold Groznov

Supervisors: Prof. Dr. Thomas März, Prof. Dr. Elke Hergenröther
Darmstadt University of Applied Sciences

Motivation

Fluorescence microscopy is an established tool in life sciences, used to visualize biological structures and processes. This imaging technique works by illuminating the specimen and thus exciting fluorophores located within the specimen with light of specific wavelengths, causing them to emit light that can be detected and used to generate detailed images. However, it requires balancing key imaging parameters, including speed, spatial resolution, light exposure, and imaging depth. Major challenges in fluorescence microscopy are photobleaching and phototoxicity, which necessitate minimizing photon exposure by reducing the duration or intensity of light. This, in turn, degrades image quality, highlighting the importance of computational methods for improving image reconstruction.

Typical distortions in fluorescence microscopy images are related to the imaging process itself. The following origins of degradation can be identified:

- Blur: described by the Point Spread Function (PSF) - a function mapping any light-emitting point in the specimen to a larger volume in the image. PSF depends on the emission wavelength and the properties of the objective lens [1].
- Noise of two primary types:
  - Photon noise, caused by the quantum nature of light. It follows a Poisson distribution and is signal-dependent, i.e. its standard deviation increases with higher signal levels.
  - Read noise, also known as electronic noise, which depends on the precision of the detector in a microscopic system. It is signal-independent and typically follows a Gaussian distribution with a mean of zero and a constant standard deviation [1].

Recent advances in machine learning have introduced CNN-based methods, showing promise in reconstruction of fluorescence microscopy images [4]. However, training these models from scratch can be impractical due to limited microscopy datasets and high computational costs. A solution is to leverage pre-trained networks, initially trained on large datasets of generic images with synthetic Gaussian noise. These models can then be fine-tuned with fluorescence microscopy data, enabling more effective image reconstruction while reducing the need for extensive training data.

The pre-training step utilizes large datasets consisting of regular photographs, which are significantly easier to obtain compared to the relatively scarce microscopy images. This approach leverages the benefits of prior training, since the computational load for model pre-training has already been incurred in a different study and is therefore amortized.

The focus is restricted to a specific class of models: Gaussian denoising networks. These models are well-studied and their fine-tuning has demonstrated competitive performance in reconstructing computed tomography (CT) image data [3]. Additionally, Gaussian denoisers provide a strong foundation for further improvements, as they constitute an active field of research.

Research Questions

To investigate the potential of pre-trained Gaussian denoiser networks for fluorescence microscopy image reconstruction, the following research questions are formulated:

- Can the use of pre-trained Gaussian denoisers be beneficial for fluorescence microscopy image reconstruction?
- How do different Gaussian denoising network architectures perform when fine-tuned for fluorescence microscopy image reconstruction?
- Which optimization techniques have the most significant impact during the fine-tuning process of Gaussian denoisers?

Fluorescence Microscopy Datasets

Several publicly available fluorescence microscopy datasets were considered for model training, with the dataset published by Hagen et al. [2] being selected. This choice is based on its high-quality 32-bit images, the inclusion of data from both widefield (eight datasets) and confocal (two datasets) microscopes, and the representation of diverse biological structures such as actin, mitochondria, nucleus, and membrane. Additionally, the datasets are pre-organized into training and test images and provide test image outputs processed with three different reconstruction methods, which can serve as a direct baseline for comparison.

Each of the ten datasets encompasses 76 to 104 image pairs. The datasets consist of paired images. In each pair, a raw image is obtained either with shorter light exposure (in case of widefield microscopy) or with reduced laser power (in case of confocal microscopy). The size of the images varies from 512x512 to 2048x2048 pixels. Fig. 1 gives an example of an image pair drawn from one of the datasets.

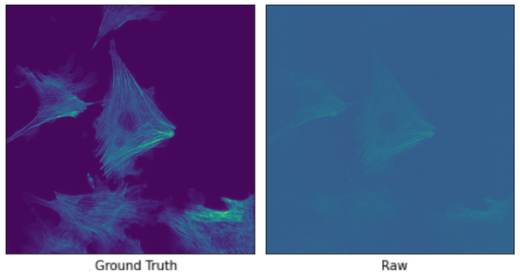

Figure 1. Example of an image pair from the dataset representing actin, obtained with light exposure of 1000 ms (ground truth) and 15 ms (raw image).

Gaussian Denoising Networks

Three pre-trained Gaussian denoising networks were selected based on several key criteria. These models feature recently introduced architectures and have demonstrated state-of-the-art performance in image denoising. Additionally, a pre-trained variant capable of processing one-channel gray-scale images was available. The selected networks represent diverse architectural designs, ensuring a broad evaluation. The following three networks have been chosen based on these considerations:

1. **DRUNet** [6] (2021): a UNet-based architecture with additional residual connections.
2. **Restormer** [5] (2021), employing a modified multi-head attention module that in contrast to the conventional self-attention mechanism has linear complexity.
3. **KBNet** [7] (2023) with kernel bases attention module as its essential ingredient. The module encompasses a series of learnable kernel bases which are fused linearly.

Experimental Setups

Three pre-trained denoising models are trained on ten datasets, and image reconstruction output is evaluated with the metrics Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). Hyperparameter tuning is applied in some cases. To compare with pre-trained models, additional experiments are conducted with denoiser models trained from scratch. Furthermore, data augmentation with rotation is applied in select cases to assess its impact on performance.

Evaluation of Results

For the given datasets, the pre-trained denoiser networks show slightly better performance in comparison to the Content Aware Restoration (CARE) method [4] across many metrics, while achieving comparable performance on the remaining ones. For the datasets containing the large images, the performance of the transformer-based networks is less impressive, which may be due to the inability to conduct a proper hyperparameter optimization, since their training requires high computational resources. Table 1 resumes the performance of the pre-trained denoiser networks, trained on one of the datasets.

Model	PSNR	SSIM	Training Duration
CARE	36.94	0.950	-
DRUNet	37.06	0.925	$\sim$ 1.6 hours
Restormer	37.56	0.926	$\sim$ 13 hours
KBNet	37.73	0.926	$\sim$ 4.3 hours

Table 1. Performance of the pre-trained denoiser networks, trained on the dataset representing nucleus, compared to the results of CARE method [4].

Considering the comparison with networks trained from scratch, the results show a significant performance gap at the early stages of training, with pre-trained networks demonstrating much higher initial performance. While networks trained from scratch tend to catch up over time, this process requires more epochs to achieve comparable results. The results reveal no substantial improvement in performance achieved through data augmentation with rotation.

Conclusion

1. Pre-trained Gaussian denoisers can enhance fluorescence microscopy image reconstruction. These models, originally developed for general denoising tasks, are straightforward to adapt for this specialized domain and capable of achieving state-of-the-art performance. In some cases, they even exceed the performance of methods specifically designed for microscopy image reconstruction.
2. UNet- and transformer-based architectures are both effective for fluorescence microscopy image reconstruction. While transformers may offer slightly better results, their high computational cost limit the possibility of hyperparameter tuning. In contrast, the investigated UNet-based model offered a less resource-intensive alternative while still achieving competitive performance, making this architecture preferable when computational resources or training speed are constraints.
3. Hyperparameter selection plays a crucial role in fine-tuning Gaussian denoisers. Mean Absolute Error (MAE) proved to be a simple yet effective loss function for this task, while SSIM-based loss offers some improvements but adds complexity. The learning rate also emerged as a key hyperparameter. Careful tuning of these parameters is essential to maximize the performance of fine-tuned models.

References

[1] Peter Bankhead. Analyzing fluorescence microscopy images with imagej, 2014. URL `https://petebankhead.gitbooks.io/imagej-intro/content/`.

[2] Guy M Hagen et al. Fluorescence microscopy datasets for training deep neural networks. *GigaScience*, 10(5):giab032, 05 2021. ISSN 2047-217X. doi: 10.1093/gigascience/giab032.

[3] Tim Selig et al. Enhanced low-dose ct image reconstruction by domain and task shifting gaussian denoisers, 2024. URL `https://arxiv.org/abs/2403.03551`.

[4] Martin Weigert et al. Content-aware image restoration: Pushing the limits of fluorescence microscopy. *bioRxiv*, 2018. doi: 10.1101/236463.

[5] Syed Waqas Zamir et al. Restormer: Efficient transformer for high-resolution image restoration, 2022. URL `https://arxiv.org/abs/2111.09881`.

[6] Kai Zhang et al. Plug-and-play image restoration with deep denoiser prior, 2021. URL `https://arxiv.org/abs/2008.13751`.

[7] Yi Zhang et al. Kbnet: Kernel basis network for image restoration, 2023. URL `https://arxiv.org/abs/2303.02881`.