Comparing Deep Learning And Traditional Denoising Methods For Voltage Imaging

Jan Willem Eriks[1]

Supervisor(s): Nergis Tömen[1]**, Alejandro Castañeda Garcia**[1]

[1]EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 22, 2025

Name of the student: Jan Willem Eriks
Final project course: CSE3000 Research Project
Thesis committee: Nergis Tömen, Alejandro Castañeda Garcia, Chirag Raman

***Abstract*—Voltage imaging enables high resolution recordings of neuronal activity but suffers from low signal-to-noise ratios (SNR), primarily due to photon shot noise. Traditional denoising methods like VST-GAT and Penalized Matrix Decomposition (PMD) offer effective noise reduction but often trade off temporal and spatial resolution. Recently, deep learning-based denoising methods, such as CellMincer, have emerged as promising alternatives due to their ability to learn complex signal models without requiring clean training data. This paper compares the performance of traditional and deep learning methods for denoising voltage imaging data using both synthetic and in vivo datasets. Metrics such as SNR, PSNR, and tSNR were used to evaluate performance. The results show that CellMincer outperforms traditional methods on synthetic data and performs competitively on real in vivo recordings, suggesting the viability of self-supervised deep learning for voltage imaging denoising. PMD remains a strong baseline with robust performance across datasets. This comparative study highlights both the potential and current limitations of deep learning approaches and suggests directions for future improvement.**

***Index Terms*—Voltage imaging, denoising, deep learning, signal-to-noise ratio, CellMincer, PMD, VST-GAT, BM3D**

I. Introduction

THE brain is one of the most complex biological systems of the body, and while the physical and cellular properties are now well documented, the inner workings of the machine that controls us is still a large question mark. The last years our ability to observe the brain has made exciting progressions. By using positron emission tomography (PET) it was possible to map regions of brain activation by measuring glucose consumption [9]. Later, the resolution of measuring brain activity by region was increased by using magnetic resonance imaging (MRI) [2]. These techniques allow for the functional mapping of the brain, however they are constrained by their low spatial resolution. Both PET and MRI can only measure the average neuron activation for a given area, which may consist of many thousands of neurons. To observe the much smaller individual neuron circuitry a more sensitive method is required. While it is possible to use an array of electrodes to measure individual neuron activation such setups are unpractical because of their size and placement [10]. The use of fluorescence voltage indicators was first demonstrated in 1968 [20] and the ability to measure multiple neurons simultaneously has first been pioneered in 1977 [19]. Since 1977 the most significant improvements to voltage imaging have been in photo sensors and the Genetic Encoded Voltage Indicators (GEVI's). GEVI's are fluorescent proteins that change their level of photon emission based on the voltage. Camera sensors, which are commonly available, have made revolutionizing improvements in resolution and ability to capture at high speed required for voltage imaging [7].

The largest contributor of noise for fluorescence voltage imaging as discussed in [19] is the so called shot noise. This comes from the fact that when a light source is sampled by only measuring a low number of photons the resulting signal will represent a Poisson distribution. The most common ways of eliminating shot noise often involves temporal and spatial averaging. This is not suitable for voltage microscopy because the loss in temporal and spatial resolution makes the action potential signal extraction of the neurons less accurate [13].

Therefore, different noise suppression methods that boost the signal-to-noise ratio (SNR) are required that better preserve the spatial and temporal resolution. Over the years multiple algorithms have been developed that have shown very good results. More recently deep learning methods have claimed to achieve even higher SNR's. This paper tries to explain the different approaches taken by traditional methods and the new deep learning based ones. This paper will also perform a comparative study to show how well these methods generalize and wether they achieve their claimed results. This goal is achieved by answering the main research question for this paper:

> *What deep learning-based denoising methods can be effectively applied to microscopy and voltage imaging, and how do they compare to traditional techniques?*

To answer the main question, the following sub-questions will be addressed:

1) *What traditional denoising methods are used for denoising voltage imaging?*
2) *What deep learning-based denoising methods can be used for denoising of voltage imaging?*
3) *How do deep learning-based and traditional methods perform in denoising voltage imaging data, as measured by improvements in signal-to-noise ratio?*

This papers main contributions are:

- Comparison between the performances and evaluation of PMD, VST-GAT and CellMincer.
- Evaluation of tSNR as a metric for denoising methods.
- Comparison between real and synthetic voltage imaging data.

II. Background

Voltage imaging data is unique due to the presence of structured noise and fast, spatially correlated, activity patterns. As a result, general-purpose denoising techniques may not perform optimally. The low SNR characteristic of voltage imaging can obscure biologically meaningful signals, complicating downstream analyses. This section provides an overview of the main challenges and recent advances in denoising voltage imaging data.

A. Sources of Noise in Voltage Imaging

In voltage imaging, noise arises from several sources, including:

- Photon shot noise due to the stochastic nature of photon emission and detection
- Instrumental noise from the imaging hardware (e.g., detectors, scanners)
- Biological variability and motion artifacts

These noise sources result in fluctuations that can be difficult to separate from the true neural signal, motivating the need for robust denoising methods.

B. Existing Denoising Approaches

Several approaches have been used to improve the quality of voltage imaging data:

a) Traditional methods: Traditional denoising methods for voltage imaging typically rely on statistical assumptions about the noise and signal. Common techniques include variance-stabilizing transformations (VST), wavelet filtering, and Gaussian smoothing. These methods are lightweight and interpretable, but are limited by their reliance on hand-crafted models that may not generalize well across datasets with different noise or signal characteristics.

b) Deep learning methods: Deep learning methods have high representational capacity and can learn arbitrary signal models [11]. This makes them particularly effective for modeling complex, high dimensional data. The ability to learn the signal model themselves can be a major advantage over traditional methods as deep learning models do not require the creation of statistical models specifically tuned for voltage imaging.

However, until recent advances in self-supervised learning, using such models for the purpose of denoising voltage imaging remained a challenge due to the absence of clean training data. Self-supervised models exploit statistical redundancies in the data itself to learn signal-preserving denoisers. In particular, self-supervised approaches such as Noise2Noise [15] first demonstrated this ability.

III. Methodology

This section will explain the setup of the experiment. It will give an explanation of the algorithms considered and why they were chosen. Equally as important this section will explain the datasets and metrics that are used to evaluate the denoising algorithms.

A. Algorithms

Due to time constrains not all traditional and deep learning methods could be included in this paper. therefore a selection was made. For traditional methods VST-GAT and PMD were selected because of their relevance in the field of voltage imaging and their claimed performance. For deep learning methods CellMincer was selected. Cellmincer was selected because it was only recently published and claims higher denoising performance of very prominent denoising algorithms such as SUPPORT [12] and DeepCad-RT [16]. A more detailed explanation of the methods is given below:

- **VST-GAT: Variance Stabilizing Transformation with the Generalized Anscombe Transformation** [17]

Variance stabilization has commonly been used for denoising and it lends itself especially well for the denoising of voltage imaging. Because the noise is mostly signal-dependent, the observed photons from the GEVI are modeled with a Poisson distribution. This gives us the following properties of the observed signal Y: $Y \sim Poisson(\lambda)$, $\mathbb{E}[Y] = \lambda$, and $\text{Var}(Y) = \lambda$. It breaks with the common convention assumed about noise, namely that it is constant. This means brighter regions will have more noise than dimmer regions. VST attempts to decouple the noise dependence on the signal, thereby giving it a constant variance that can be modeled to a Gaussian distribution [22] [17]. VST-GAT uses the Generalized Anscombe Transformation for variance stabilization.

This has the benefit that denoising Gaussian noise is a very well studied topic with many good performing algorithms to choose from. In [17] multiple Gaussian denoising methods were evaluated. They found that BM3D [4] achieved the highest SNR when denoising images of fluorescent cells. Fluorescent cells imaging exhibit many of the same challenges of voltage imaging, mostly that the signal follows the Poisson distribution. Therefore the combination of VST-GAT with BM3D was used, which will be called VST from this point onward.

- **PMD: Penalized Matrix Decomposition** [3]

PMD has been the latest notable traditional algorithm for voltage imaging developed. Even though PMD is a traditional method, newly developed deep learning still compare their denoising performance to PMD (SUPPORT, DeepInterpolation, CellMincer) [6] [14] [21]. PMD has achieved remarkable denoising performance and is specifically designed for voltage imaging. PMD works by taking advantage of three key observations about voltage imaging [3]:

1) The signal sources are spatially local.
2) The signal is structured both temporally and spatially, whereas noise is temporally and spatially uncorrelated.
3) The signal is low-rank.

This statistical model of the signal and noise allows for a decomposition of the original matrix $\mathbf{Y}$ of shape $d \times T$, where d is the number of pixels in a frame and T is the number of frames. PMD converts $\mathbf{Y}$ into a low rank approximation $\hat{\mathbf{Y}} = \mathbf{UV}$ where $\mathbf{U}$ contains the spatial information of the signal and $\mathbf{V}$ contains the temporal information e.g. neuron activation. $\mathbf{U}$ and $\mathbf{V}$ are iteratively optimized so that the residual not captured information by $\mathbf{U}$ and $\mathbf{V}$ represents unstructured noise. This comes with the important assumption that neurons stay spatially static in the movie, which does not hold for all the data as seen in subsection III-B.

- **CellMincer: A U-Net based deep learning model** [21]

CellMincer is a relatively new proposed method for denoising voltage imaging. The model architecture consists of two main components: a spatial feature extractor based on a modified U-Net and a temporal post-processor that performs pixel-wise denoising over time.

The U-Net, originally introduced in [18], is used here as a frame wise encoder decoder network. It operates on individual frames to extract high resolution spatial features through a contracting path (encoder) and reconstructs spatially detailed embeddings via an expanding path (decoder). In CellMincer, these embeddings represent the spatial structure of the neuronal activity in each frame. In Figure 1 is a visualization given of how the U-Net performs the contracting and expanding. The green layer indicates the global feature map which is extracted in a pre-processing step and contains spatial temporal information of the pixels.

The output embeddings from a short sequence of consecutive frames are then fed into a 1D temporal convolutional module, which processes the time series of embeddings for each pixel independently. This temporal

post-processor predicts the denoised value for the center frame in the sequence, leveraging temporal correlations while preserving unique frame specific signals.

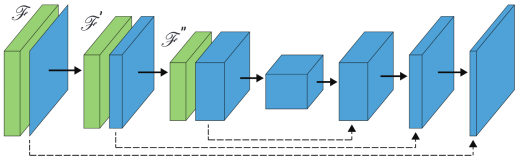


Fig. 1: A visualization of CellMincer's U-Net, where blue indicates a frame and green indicates the global feature map. Taken from [21]

.

For the purpose of this research a pre-trained model of CellMincer was used. This model was trained on data similar to that of the Optosynth dataset, the effects of this on the results is discussed in section VI.

B. Datasets

For this paper two datasets were selected: HPC2 and Optosynth. HPC2 is a real in vivo dataset of voltage imaging and Optosynth is a synthetic dataset that is generated by the Optosynth framework. Real in vivo voltage imaging data is not abundantly available as the data requires a complex setup to be collected [1]. Therefore synthetic data is used in this paper to increase the ability to evaluate the methods. Synthetic data also has the benefit of having a ground truth available. With the ground truth much more established evaluation metrics become available, these are discussed more in the Evaluation section. A more detailed description of the datasets is given below:

- **HPC2** [1]

The HPC2 dataset was collected from the hippocampus of a mouse. As common with voltage imaging the mouse was put in fixed position where through a small hole in its cranium a laser illuminates the GEVI. The HPC2 dataset is a collection of movies with a 498 x 116 pixel resolution, where each movie has 15,000 frames. A sample frame of the dataset is displayed in Figure 2. For this paper two movies were selected due to the large size of the total dataset. In section VI the effects on the results of using only a subset of the dataset are discussed.

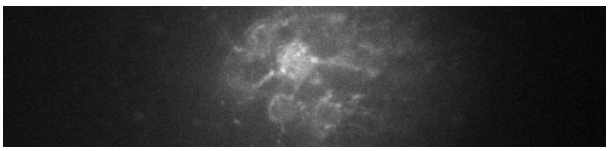


Fig. 2: A sample of from the HPC2 dataset

- **Optosynth** [21]

Synthetic voltage imaging tries to be as closely resembling to in vivo voltage imaging as possible. This means modeling the physical properties of neurons such as their action potential propagation and spatial distribution in the brain. Optosynth uses the Allen Brain Atlas, a database containing properties of brain areas [8], to create a realistic simulation of a mouse primary visual cortex. The Optosynth dataset contains multiple movies that are 512 x 180 pixels with each sample containing 7,000 frames. The dataset contains multiple samples of each movie where each sample contains a different level of noise. For this paper only the samples with the highest noise level were considered. The denoising performance of the methods on lower levels of noise is not within the topic of this research. This decision was made to preserve time and denoising low SNR movies is much more of a challenge. The effects of this decision on the results are discussed in section VI.

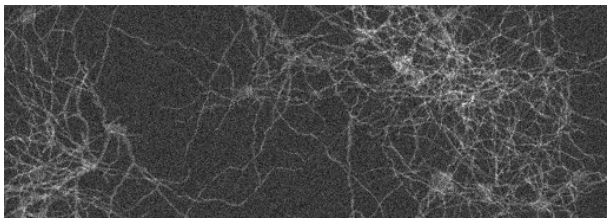


Fig. 3: A sample from the Optosynth dataset

Some notable differences between the HPC2 and Optosynth datasets are the amount of neurons a single frame contains. The Optosynth has a much wider field of view Figure 3 and shows 485 neurons [21] while HPC2 only shows a couple of neurons. Another key difference is that HPC2 contains some significant motion artifacts and sometimes the camera loses focus. Optosynth does not simulate these phenomena, however they appear very common in voltage imaging.

C. Evaluation

To evaluate the performance of the denoising algorithms the whole denoised image was evaluated. Other papers sometimes analyze the neuron activation traces after denoising. While they are closely related and better image denoising should result in better neuron activation traces, it is worth to mention the difference. The following metrics were used to measure the performance of the algorithms:

- SNR (Signal-to-Noise Ratio)
- PSNR (Peak Signal-to-Noise Ratio)
- tSNR (Temporal Signal-to-Noise Ratio)

All these metrics measure a form of a signal to noise ratio. Therefore, a higher ratio indicates a greater performance of the denoising algorithm.

- **SNR and PSNR**

SNR and PSNR are the most commonly used metrics for assessing image quality and denoising performance. SNR and PSNR measure how close the denoised output is to the clean ground truth. The main difference between the two is that PSNR is insensitive to the brightness of the signal. For images of high average pixel brightness high levels of noise can still report high SNR. SNR and PSNR can only be measured on the syntethic dataset and not the in vivo becuase of the need for a ground truth.

SNR and PSNR are defined as:

$$\text{SNR} = 10\log_{10}\left(\frac{\sum_{i=1}^{N} y_i^2}{\sum_{i=1}^{N}(x_i - y_i)^2}\right)$$

$$\text{PSNR} = 10\log_{10}\left(\frac{MAX^2}{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2}\right)$$

Where x_i is the denoised output, y_i is the clean ground truth, and N is the total number of pixels in the movie. *MAX* denotes the maximum possible pixel value.

- **tSNR**

For real in vivo data SNR and PSNR cannot be calculated due to the absence of a ground truth, therefore tSNR is used as substitute metric. tSNR is commonly used as an MRI performance metric [5] and indicates the stability of the signal. A higher tSNR indicates a more temporally stable signal. Here the assumption is made that low SNR data contains an unstable signal and high SNR data a stable signal.

tSNR is defined as:

$$\text{tSNR} = 10\log_{10}\left(\frac{\mu}{\sigma}\right)$$

Where μ is the temporal mean of the signal at a given pixel across frames, and σ is the corresponding standard deviation. To allow for evenly scaled metrics tSNR is calculated in decibels.

This combination of metrics allows us to evaluate both absolute performance on synthetic data and relative improvements in stability on the in vivo data, providing a comprehensive assessment of the denoising methods.

IV. Results

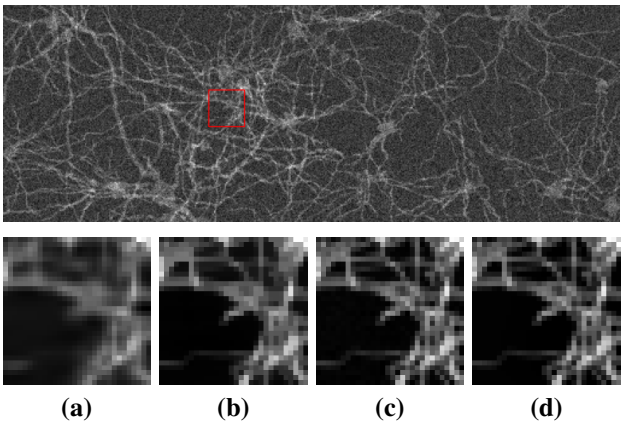
This section will give an overview of the performance of the denoising methods. Table I and Table II show the results for the performed methods together with the baseline performance of the original noisy data for comparison. Table I shows the results for the synthetic dataset, and Table II for the in vivo dataset.

TABLE I: Quantitative performance on the Optosynth dataset

Method	SNR	PSNR	tSNR
VST	30.74	66.87	29.59
PMD	49.48	85.6	31.47
CellMincer	58.28	94.39	29.63
Original	33.12	69.23	16.63

For the Optosynth dataset, CellMincer performed significantly better at the SNR and PSNR metrics. It outperformed the other metrics on SNR and PSNR which are considered the most reliable metrics. PMD did perform better on tSNR than CellMincer indicating a more stable signal. Overall, VST performed the worst all the methods.

Figure 4 shows a sample frame of the original noisy image. To visualize the different results between the denoising methods a zoomed in section marked by the red rectangle of their output of that frame is shown below. The higher pixel values indicate excitement of the neuron. Figure 4 shows that CellMincer (c) and PMD (b) have a very close resemblance to the ground truth. VST (a) did struggle more with the high levels of detail as the resulting image is quite blurry, which corresponds to the low SNR and PSNR values.



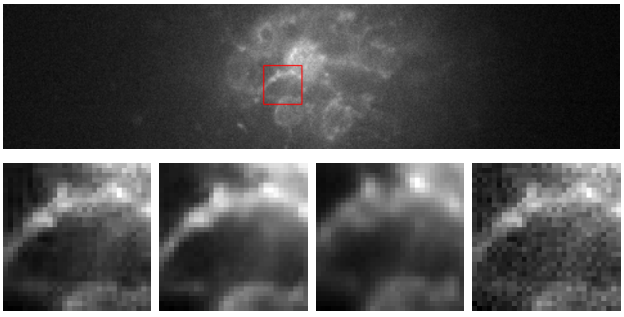
(a) **(b)** **(c)** **(d)**

Fig. 4: Comparison of different methods: (a) VST, (b) PMD, (c) CellMincer and (d) the ground truth.

TABLE II: Quantitative performance on the HPC2 dataset

Method	tSNR
VST	19.96
PMD	20.68
CellMincer	21.0
Original	13.98

For the HPC2 dataset, only the tSNR could be calculated because of the absence of the ground truth. CellMincer did achieve the highest tSNR score. However as shown in Figure 5, PMD (b) did result in a sharper image with more details retained. VST (a) again performed the worst of the three methods. In contrast to the Optosynth dataset VST did not result in a blurry image. However, Figure 5 shows that VST did struggle to deal with the high noise level of the HPC2 dateset. The resulting sample still has some noise left where PMD and CellMincer (c) achieved a smoother image.



(a) **(b)** **(c)** **(d)**

Fig. 5: Comparison of different methods: (a) VST, (b) PMD, (c) CellMincer, (d) the noisy original.

V. Responsible Research

For this paper, only publicly accessible data and code were used. Almost all papers publicate their code and often also their models. This has been essential this research as re-implementing and training the selected methods would require more time then is available. While the code is often easily available, the datasets are much more locked down. Many papers will share their data only after a signed contract confirms the intent and authenticity of the requester. While this is understandable, it does limit how easy a study can be reproduced. Due to time constrains and the previously stated issue only freely accessible datasets were used in this paper. Furthermore, to make the results of this paper as trustworthy as possible all code has been made publicly available[1], together with the code to generate the figures.

VI. Conclusion & Discussion

The main goal of this research was to identify the differences in performance between traditional and deep learning methods for denoising voltage imaging data. The results show that deep learning models can outperform traditional statistical models, even when these deep learning models are only trained on noisy synthetic data. Below a more in depth analysis of the results and findings of this paper can be found.

A. Methods

The fact that CellMincer showed strong results on the Optosynth dataset, but performed worse on the HPC2 dataset was expected. The model was trained on data similar to that of Optosynth, which explains its better performance there. On the HPC2 dataset, CellMincer did manage to remove a significant amount of noise, but it may have done so at the cost of image sharpness as seen in Figure 5. Nevertheless, its ability to generalize to previously unseen real in vivo data demonstrates the viability of this method. It also suggests that further improvements are possible if the model is trained directly on real in vivo data.

PMD performed consistently well on both the Optosynth and HPC2 datasets. Although it scored lower on SNR and PSNR compared to CellMincer, its results were still impressive. PMD's ability to preserve sharp edges while handling high levels of noise in real in vivo data even surpassed that of CellMincer. It is clear why PMD is still often used as a baseline for comparing new methods since its performance remains competitive. PMD does not require a training phase or extensive hyperparameter tuning, making it easy to apply. This means it can be used without a deep understanding of its internal mechanics and still produce high-SNR voltage imaging data.

VST performed the worst of all the methods. The fact that VST performed worse than newer methods like PMD and CellMincer is not surprising. However, the only metric where it out performed the original noisy data is tSNR. As can be seen in Table I it even degraded the SNR and PSNR scores compared to the original noisy data. This means it is not a suitable technique for denoising voltage imaging. The very bad performance can be explained by the fact that it was originally not specifically designed for denoising voltage imaging. In the research that introduced VST, BM3D was found to be the best performing method on pictures of fluorescent cells. Apparently this did not generalize well to movies of voltage imaging. Another factor that can explain its lack in performance is that BM3D only processes one frame at the time. Therefore it did not take advantage of temporal structures in the data, while both CellMincer and PMD extensively use temporal information.

B. Datasets

The Optosynth dataset provides a useful and flexible synthetic benchmark for evaluating denoising algorithms. Unlike real in vivo data, it includes clean ground truth, which enables direct quantitative comparisons using metrics such as SNR and PSNR. However, it lacks several real-world properties, such as tissue movement, neuron displacement, and motion blur. These aspects are common in in vivo recordings and pose significant challenges for denoising methods. Despite these limitations, the utility of Optosynth remains substantial. The ability to generate large volumes of data under controlled conditions, with adjustable noise levels and a known ground truth, makes it an indispensable tool for benchmarking and development. This paper finds that the flexibility and reproducibility of Optosynth dataset offsets its lack of full biological realism. Optosynth allows for systematic and interpretable performance evaluations that are not possible with real datasets alone.

As the combined complete Optosynth and HPC2 datasets contain over 200,000 frames, only a subset was used for evaluation due to time and computational constraints. This selective sampling is unlikely to have introduced significant bias, as frame-to-frame variation within each dataset was found to be relatively minor and consistent.

C. Metrics

For this research, SNR, PSNR, and tSNR were used as evaluation metrics. When assessing denoising performance with access to ground truth, SNR and PSNR are widely regarded as standard metrics. Their individual usefulness is well established and largely undisputed. PSNR alone is often considered sufficient for evaluating image quality, while SNR remains valuable due to its broad adoption, which facilitates comparison across studies. A known limitation of SNR is its sensitivity to absolute pixel intensity, which can lead to inflated values in brighter images. However, this issue does not apply in this study, as all images used had comparable maximum pixel values. tSNR was employed as an alternative metric for in vivo data, where ground truth is not available. While it is a convenient and widely used measure of temporal signal stability, it did not show a strong correlation with SNR and PSNR. Although useful, tSNR alone was found to not fully capture the performances of denoising methods.

[1] https://github.com/JanDeDinoMan/Research-Project

D. Comparison

Direction comparisons between this research and that of Ioan Leolea is not possible due to differences between evaluation methods. As discussed in subsection III-C there are multiple ways to measure SNR, PSNR and tSNR. However, in the research of Ioan, the method AP-BSN performs the worst on the metric PSNR while performing when considering tSNR. This reinforces the idea that tSNR is not a reliable metric to measure the denoising performance.

To conclude, CellMincer performed the best on the synthetic data, with a close second on real in vivo data. Considering its current performance and the fact that CellMincer was only trained on synthetic data means its performance is impressive. This performance also shows the viability of deep learning methods for denoising voltage imaging data. Therefore CellMincer is concluded to be the best performing denoising method researched in this paper.

VII. Future recommendations

In this paper three points were identified which could be further researched:

- The CellMincer model used in this study was trained solely on synthetic data. To fully evaluate its potential, CellMincer should be trained on in vivo data similar to the datasets on which it is tested. We expect this could lead to significant improvements in performance.
- VST showed poor performance on voltage imaging data, though this may be partly due to the choice of BM3D as the denoising algorithm. Future work could explore replacing BM3D with a method more suitable for voltage imaging. Particularly one that incorporates temporal information which may yield considerably better results.
- While Optosynth provides a valuable synthetic benchmark, it does not simulate several key physical phenomena that contribute to noise in real voltage imaging, such as fluid motion and sensor instability. Extending Optosynth to model these effects could significantly enhance its utility as a dataset for evaluating denoising methods.

References

[1] Yosuke Bando, Ramdas Pillai, Atsushi Kajita, Farhan Abdul Hakeem, Yves Quemener, Hua-an Tseng, Kiryl D. Piatkevich, Changyang Linghu, Xue Han, and Edward S. Boyden. Real-time Neuron Segmentation for Voltage Imaging. In *2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 813–818, December 2023. arXiv:2403.16438 [eess].

[2] J. W. Belliveau, D. N. Kennedy, R. C. McKinstry, B. R. Buchbinder, R. M. Weisskoff, M. S. Cohen, J. M. Vevea, T. J. Brady, and B. R. Rosen. Functional Mapping of the Human Visual Cortex by Magnetic Resonance Imaging. *Science*, 254(5032):716–719, November 1991. Publisher: American Association for the Advancement of Science.

[3] E. Kelly Buchanan, Ian Kinsella, Ding Zhou, Rong Zhu, Pengcheng Zhou, Felipe Gerhard, John Ferrante, Ying Ma, Sharon Kim, Mohammed Shaik, Yajie Liang, Rongwen Lu, Jacob Reimer, Paul Fahey, Taliah Muhammad, Graham Dempsey, Elizabeth Hillman, Na Ji, Andreas Tolias, and Liam Paninski. Penalized matrix decomposition for denoising, compression, and improved demixing of functional imaging data, January 2019. Pages: 334706 Section: New Results.

[4] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, August 2007.

[5] Beatriz Dionisio-Parra, Florian Wiesinger, Philipp G. Sämann, Michael Czisch, and Ana Beatriz Solana. Looping Star fMRI in Cognitive Tasks and Resting State. *Journal of Magnetic Resonance Imaging*, 52(3):739–751, 2020. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jmri.27073.

[6] Minho Eom, Seungjae Han, Pojeong Park, Gyuri Kim, Eun-Seo Cho, Jueun Sim, Kang-Han Lee, Seonghoon Kim, He Tian, Urs L. Böhm, Eric Lowet, Hua-an Tseng, Jieun Choi, Stephani Edwina Lucia, Seung Hyun Ryu, Márton Rózsa, Sunghoe Chang, Pilhan Kim, Xue Han, Kiryl D. Piatkevich, Myunghwan Choi, Cheol-Hee Kim, Adam E. Cohen, Jae-Byum Chang, and Young-Gyu Yoon. Statistically unbiased prediction enables accurate denoising of voltage imaging data. *Nature Methods*, 20(10):1581–1592, October 2023. Publisher: Nature Publishing Group.

[7] Takeharu G. Etoh and Quang A. Nguyen. Evolution of High-Speed Image Sensors. In Kinko Tsuji, editor, *The Micro-World Observed by Ultra High-Speed Cameras: We See What You Don't See*, pages 81–101. Springer International Publishing, Cham, 2018.

[8] Allen Institute for Brain Science. Brain Map - brain-map.org.

[9] P. T. Fox, M. E. Raichle, M. A. Mintun, and C. Dence. Nonoxidative glucose consumption during focal physiologic neural activity. *Science (New York, N.Y.)*, 241(4864):462–464, July 1988.

[10] Amiram Grinvald, D. Shoham, A. Shmuel, D. Glaser, I. Vanzetta, E. Shtoyerman, H. Slovin, C. Wijnbergen, R. Hildesheim, and A. Arieli. In-vivo Optical Imaging of Cortical Architecture and Dynamics. In Uwe Windhorst and Håkan Johansson, editors, *Modern Techniques in Neuroscience Research*, pages 893–969. Springer, Berlin, Heidelberg, 1999.

[11] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366, January 1989.

[12] Kendrick Kay. The risk of bias in denoising methods: Examples from neuroimaging. *PLOS ONE*, 17(7):e0270895, July 2022. Publisher: Public Library of Science.

[13] Rishikesh U. Kulkarni and Evan W. Miller. Voltage Imaging: Pitfalls and Potential. *Biochemistry*, 56(39):5171–5177, October 2017.

[14] Jérôme Lecoq, Michael Oliver, Joshua H. Siegle, Natalia Orlova, Peter Ledochowitsch, and Christof Koch. Removing independent noise in systems neuroscience data using DeepInterpolation. *Nature Methods*, 18(11):1401–1408, November 2021. Publisher: Nature Publishing Group.

[15] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2Noise: Learning Image Restoration without Clean Data, October 2018. arXiv:1803.04189 [cs].

[16] Xinyang Li, Yixin Li, Yiliang Zhou, Jiamin Wu, Zhifeng Zhao, Jiaqi Fan, Fei Deng, Zhaofa Wu, Guihua Xiao, Jing He, Yuanlong Zhang, Guoxun Zhang, Xiaowan Hu, Xingye Chen, Yi Zhang, Hui Qiao, Hao Xie, Yulong Li, Haoqian Wang, Lu Fang, and Qionghai Dai. Real-time denoising enables high-sensitivity fluorescence time-lapse imaging beyond the shot-noise limit. *Nature Biotechnology*, 41(2):282–292, February 2023. Publisher: Nature Publishing Group.

[17] Markku Makitalo and Alessandro Foi. Optimal Inversion of the Generalized Anscombe Transformation for Poisson-Gaussian Noise. *IEEE Transactions on Image Processing*, 22(1):91–103, January 2013.

[18] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation, May 2015. arXiv:1505.04597 [cs].

[19] B. M. Salzberg, A. Grinvald, L. B. Cohen, H. V. Davila, and W. N. Ross. Optical recording of neuronal activity in an invertebrate central nervous system: simultaneous monitoring of several neurons. *Journal of Neurophysiology*, 40(6):1281–1291, November 1977. Publisher: American Physiological Society.

[20] I Tasaki, A Watanabe, R Sandlin, and L Carnay. Changes in fluorescence, turbidity, and birefringence associated with nerve excitation. *Proceedings of the National Academy of Sciences*, 61(3):883–888, November 1968. Publisher: Proceedings of the National Academy of Sciences.

[21] Brice Wang, Tianle Ma, Theresa Chen, Trinh Nguyen, Ethan Crouse, Stephen J. Fleming, Alison S. Walker, Vera Valakh, Ralda Nehme, Evan W. Miller, Samouil L. Farhi, and Mehrtash Babadi. Robust self-supervised denoising of voltage imaging data using CellMincer. *npj Imaging*, 2(1):1–21, December 2024. Publisher: Nature Publishing Group.

[22] B. Zhang, M. J. Fadili, J.-L. Starck, and J.-C. Olivo-Marin. Multiscale Variance-Stabilizing Transform for Mixed-Poisson-Gaussian Processes and its Applications in Bioimaging. In *2007 IEEE International Conference on Image Processing*, volume 6, pages VI – 233–VI – 236, September 2007. ISSN: 2381-8549.