Denoising Microscopy Images in Voltage Imaging Videos

Overview and Feasibility of Traditional Denoising Methods

Radosław Michał Majer[1]

Supervisors: Dr. Nergis Tömen[1]**, Alejandro Castañeda García**[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 20, 2025

Name of the student: Radosław Michał Majer
Final project course: CSE3000 Research Project
Thesis committee: Dr. Nergis Tömen, Alejandro Castañeda García, Dr. Chirag Raman

Abstract

Voltage imaging is an emerging microscopy technique that can make neuroscientific research very prominent. The images obtained with this imaging method exhibit a substantial amount of noise. Currently, the new methods are developed and tested to computationally denoise voltage imaging videos with high efficiency and preservation of the video structure. This research attempted to investigate how well traditional denoising algorithms, such as different types of blur or diffusion, perform in denoising such videos. Specifically, five well-established algorithms commonly used in biomedical imaging were evaluated: Gaussian filter, bilateral filter, anisotropic diffusion, wavelet filter, and total variation minimization. The methods were applied to both real brain recordings (HPC2 dataset [1]) and synthetic videos (Broad DSP CellMincer [17]). Performance was assessed by measuring the structural similarity index (SSIM) and signal-to-noise ratio (SNR). Results suggest a trade-off between noise removal and structural preservation, with total variation minimization and anisotropic filtering performing particularly well in terms of noise suppression. These classical methods remain relevant for data exploration and visualization. For the methods to be used in a medical context, a more in-depth research should be carried out on medical data. The deep learning methods remain relevant for the high-precision applications. Code developed for the research is available online at https://github.com/Rpplctns/denoising-voltage-imaging-videos.git.

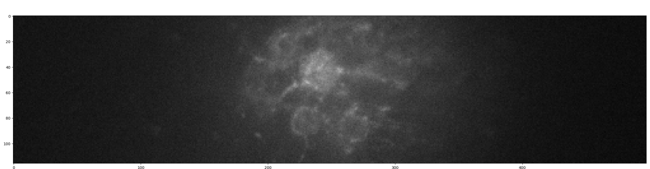

Figure 1: Example voltage imaging video frame (mouse brain)

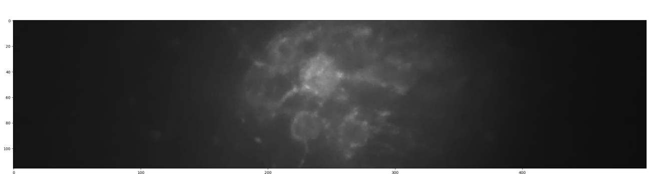

Figure 2: Example frame denoised with TV minimization

1 Introduction

Voltage imaging is a microscopy technique that has proven to be very effective in imagining the neural activity of the brain. It is a promising technique that may help the neuroscience researchers explain human behavior from a microscopic perspective of a network of single neurons. [6].

The general idea of voltage imaging is to measure the neural activity of the brain without need for use of an electrode. Use of electrode only allows to observe one neuron at the time. Voltage imaging on the other hand, could provide a bigger picture of the neural activity in a brain. For that purpose, an optical method that involves staining techniques was already proposed in the 1970s [6] [15].

The problem with the method is that the images obtained by voltage imaging often exhibit a substantial amount of noise, due to the specification of the method, described in detail in the next section. This causes huge difficulties in the output interpretation, compromising the research and causing the need for image denoising methods.

Many ways of denoising images have already been proposed in general. This includes traditional methods [2] [4] [8] and deep learning-based approaches [20]. A lot of research has been done around these methods, their efficiency is well-tested and many variants and optimizations has been proposed for different use cases. However, there is no reliable information on the feasibility of using these methods for denoising voltage microscopy images. Finally, denoising methods such as SUPPORT [5] were introduced specifically for the purpose of denoising such images. At this point, it is still a subject of investigation whether the variants of more general methods cannot be used with equally satisfactory performance.

This study aims to answer the following research question.

> **What are the traditional, non-deep learning denoising methods used for denoising biomedical and microscopy data, and how do they perform in terms of preserving signal integrity and suppressing noise in voltage imaging?**

Answering the question requires defining which traditional methods should be evaluated and describing them. Furthermore, it requires defining clear criteria of what is considered a properly denoised image, what are the evaluation criteria. Therefore, the question can be divided into the following subquestions.

> **What traditional video denoising methods are there and which ones can be used for denoising voltage imaging videos?**
>
> **How can the suitability of a denoising method be evaluated for voltage imaging?**
>
> **How do the traditional methods perform in denoising voltage imaging videos?**

This paper is structured as follows. In the Background section, the purpose of voltage imaging is presented and the need for denoising methods is described in detail. The Project Contribution section of the paper describes what has been done for the project. The Methodology section explains what datasets, methods, and assessment criteria have been used for the study and why. Finally, the Results & Deiscussion shows the results of applying evaluation metrics to denoising methods and includes the answers to the research question. The ethical aspects of the study are described in the Responsible

Research section. In the end, all the experiment process and results are summarized in the Conclusion, with the guidelines for future research.

2 Background

This section outlines the basic principles of voltage imaging. It also summarizes previous research on the method in question and denoising of multimedia.

2.1 The Idea Behind Voltage Imaging

A neural network, such as, for example, a human brain, works by transmitting electric impulses through neurons. The biological event of the nerve impulse, which is fundamental in cell-to-cell communication, is called *the action potential*. The general idea of voltage imaging is optical observation of this electrical activity. This is done by observing fluorescent indicators [15], which react for electricity. Initially, small molecules of dye served as these indicators. More modern approaches are now being used for that purpose, because of the developments in nanobiology. Today, impulses are indicated by genetically encoded voltage-sensitive proteins [6].

Among the techniques that can be used to monitor neural activity, voltage imaging is a method that allows the observation of a larger chunk of a neural network. Other approaches include electrode-based methods. However, these allow the observation of only a single neuron at the time [6]. Magnetic Resonance Imaging is a different approach that captures more than just one neuron at a time. Unlike voltage imaging, it involves observing the activity of certain regions of the brain rather than focusing on the activity of particular neurons. Another imaging technique, called calcium imaging, provides a good overview of the network. However, calcium molecules move relatively slowly, compared to the rapid nature of the action potential. Voltage imaging is therefore the best method to observe quick nuances on a larger scale.

2.2 The Opportunities in Voltage Imaging

Emerging of voltage imaging provides neuroscience with a means to observe the electrical activity of the brain with high precision and speed. It allows for the observation of a spatial neural region. It is a very powerful tool with a lot of potential for neuroscience research, which focuses on understanding how neurons connect and communicate. The technique enables observation of the brain of an awake animal or human. It can potentially lead to a deeper understanding of how electrical impulses in certain regions of the brain correlate with behavior and cognition. It can also be used to identify the response to drugs.

2.3 Challenges in Voltage Imaging

The most important problem with voltage imaging is the speed of the event. The action potential typically occurs in the span of 2ms. Its most important "rising phase" lasts about 250μs [6]. The nature of the event is relatively quick. This enforces a voltage indicator to be able to resolve changes on the tenth of millisecond time span. As a result, the image acquisition must also be fast, about 1000 Hz (compared to 1-10 Hz in calcium imaging). The little time to deliver photons leads to an image being extremely noisy. That is the reason why there is a need for computational denoising. The denoised video should possibly be used for image segmentation, or visual observation. Therefore, a video should preserve the structure of the object and possibly provide a clear, denoised view on a neuron.

2.4 Image Denoising

Many digital image processing techniques were emerging in 1960s. They focus on the analysis of mathematical properties of an image, in an attempt to restore, enhance, encode, or compress the picture [14]. Computational image restoration can be done by filtering, denoising algorithms, or machine learning. The latter has proven to be more effective in predicting the structure of the objects in the image and reconstructing the actual contents. On the other hand, the older, traditional approaches are more efficient in terms of computational resources. The traditional methods that will be discussed in this paper are introduced in the following subsections.

Gaussian Filtering

Gaussian filtering is a denoising approach, used to blur the image and remove noise in this way [8]. It involves convolving the input medium with a Gaussian kernel, which smooths the image by averaging the pixel values with their neighbors based on the weights specified by the kernel.

The method finds an application in medicine, particularly in denoising functional magnetic resonance imaging (fMRI) to remove sensor noise [9] [11].

Gaussian filtering has been chosen in this study as a classical baseline for image denoising. It is simple to implement and widely known denoising approach.

Bilateral Filtering

Bilateral filtering is a method used to denoise an image while preserving structural features, such as the edges [8]. It is similar to Gaussian filtering, but in addition to considering how close the neighboring pixels are, it also takes into account how similar they are in terms of intensity.

The method has been chosen because it balances edge preservation with noise reduction, while it is still a simple baseline method, as an improved version of Gaussian blur. Edge preservation is especially relevant for voltage images, as they contain features like neural boundaries that need to be preserved.

Anisotropic Diffusion

Unlike isotropic filters, which smooth uniformly in all directions, this approach diffuses intensity in the direction of low gradient, while preserving differences along the edges [2] [13]. The diffusion process if applied for a certain number of iterations with a certain step size. It serves similar purpose as bilateral filtering, that is, is a method used for edge-preserving denoising. However, this method works in a more complex way, through solving partial differential equations.

Anisotropic diffusion is used in medical image processing to denoise MRI scans because of its ability to preserve structural features of an image [9].

The method was chosen to be evaluated in this study as another method designed for edge preservation denoising. Its allows for a lot of flexibility and control due to the large amount of parameters.

Wavelet Denoising

Wavelet denoising decomposes the input signal using a discrete wavelet transform, and then a specified threshold is applied to filter the noise components from the signal components [4].

Wavelet denoising is commonly used in medicine and biology to denoise high-resolution microscopy images or EEG scans [16].

The algorithm has been chosen as one of the most basic signal processing methods. It can be very effective for non-stationary signals, like voltage traces in the microscopy method.

Total Variation Minimization

The total variation of an image or signal is a measure of how much the intensity changes across space or time [2]. The TV minimization method for denoising works by reducing small fluctuations, assumed to be noise, while preserving sharp changes, assumed to be signal components.

This is a powerful algorithm which is widely applied in the field. It is used to denoise MRI scans, positron emission tomography (PET) scans, and microscopy [9].

The method is a widely known denoising algorithm. It definitely should be considered in the study, as the voltage imaging photon shot noise exhibits a very high frequency.

Although many alternative denoising methods exist, this study focuses on a representative set of classic, well-established algorithms that span several key methodological categories and find use in biomedical applications. Simpler approaches such as mean or median filtering were excluded due to their poor edge preservation. Similarly, Fourier-based methods lack the localization necessary to handle rapid neural events. The selected techniques represent a diverse set of traditional approaches that are interpretable, reproducible, and widely cited in the denoising literature. This diversity allows for a systematic comparison across algorithmic families and ensures relevance to the wide range of signal characteristics present in voltage imaging data.

2.5 Recently Proposed Denoising Approaches for Voltage Imaging

Recent research around voltage imaging and deep learning resulted in the development of specific denoising techniques. These techniques allow data restoration from a voltage imaging video. In particular, a group of Korean researchers proposed a self-supervised deep learning method called SUPPORT [5]. CellMincer is an another method that uses self-supervised deep learning. CellMincer also proved to be an effective method for the use case [17]. Deep learning is extremely suitable for the task, but there remains a question if they are necessary for proper restoration of the videos.

3 Project contribution

The section provides a brief overview of the work that has been done as a part of the study and how it contributes to the field of voltage imaging postprocessing.

3.1 Denoising Methods

Through the study, five denoising approaches have been evaluated. These are listed in the Background section. Some of the methods have been implemented during the study, and for some of them the scikit and medpy implementations have been used [12] [7]. For these methods, the tweakable parameters have been enumerated in the Methodology section, and the optimal values have been selected during the experimentation process. For each method, numerical and visual results are presented as the means to value their fitness for denoising voltage imaging videos.

3.2 Evaluation

This study outlines relevant evaluation techniques. An evaluation metric is considered relevant if it provides a good overview of whether a denoising method is suitable for voltage image denoising. Apart from the quantitative metrics, qualitative results are shown and discussed in the Results section. According to the evaluation metrics, the fitness of the denoising approaches is valued and discussed.

3.3 Experiments

The paper explains the experimental setup of the study, that is, what datasets have been used, what methods have been tested, and how the parameter values were selected. The work discusses the alternatives to deep-learning approaches, and their fit for the use case, based on the results of the experiments.

4 Methodology

This section elaborates on the approach to conducting the experiments. It first outlines the datasets, then the denoising techniques that are being investigated, then the evaluation metrics, and finally a protocol for conducting the experiments.

4.1 The Datasets

In the experiments it is desirable to use data that actually resemble voltage imaging videos. However, evaluation techniques often require some ground truth to assess the fitness of a method. Because of that, in this study, two different datasets are used:

- HPC2 [1] - set of 13 voltage imaging videos of the hippocampus region of a mouse brain. The data set does not include ground truth for video denoising. Two of the images (00_02 and 00_03) are used for the evaluation of the denoising methods, and visualization of how accurate they are in preserving the contents of the video while removing the noise. While the dataset provides videos of various events happening in the brain, in this research it has to be supported by another dataset which contains a ground truth, for the experiments reasons.

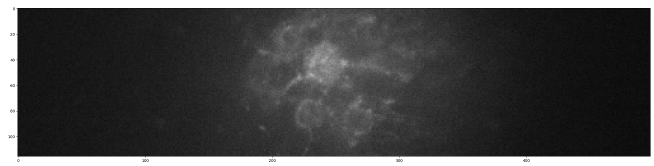

Figure 3: Example HPC2 frame

- Broad DSP CellMincer Data [17] - synthetic dataset of five videos with different noise levels. Contains a ground truth for image denoising. The dataset is created with the Optosynth tool, designed specifically for synthesizing voltage imaging videos and adding Poisson/Gaussian noise to them. It was created as a dataset for the development of the CellMincer denoising method. The study uses five videos with the highest noise available (20 units).

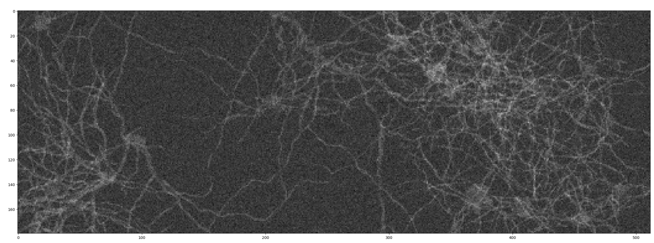

Figure 4: Example noisy frame

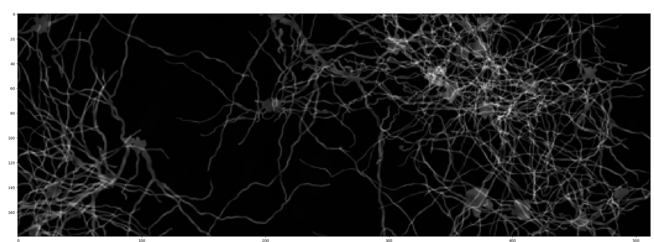

Figure 5: Example clean frame (ground truth)

The Validation Set

Hyperparameter optimization requires the validation set. The first 1000 (out of 7000) frames of the video *optosynth_1_20_5.tif* have been chosen for a validation set. The frames contain a representative piece of video. The set is sufficiently small to enable running multiple methods multiple times, with different parameter configuration.

4.2 The Denoising Approaches

This report focuses solely on traditional denoising methods. The study implements and experiments on the feasibility of these as a solution to voltage imaging problems. The following methods have been tested in the research.

Gaussian Filtering

Gaussian filtering blurs the image in order to remove the noise. In this study, it is used to denoise videos, and therefore the 3-dimensional kernel is used, with a separate variance in the Z (time) direction.

The properties of smoothing are specified by the following parameters:

- Gaussian standard deviation in X and Y (spatial) dimensions (σ_s) - varied in [0.2, 1.8], where 0 means no denoising, 1 is a standard amount used in other medical applications [11], and 1.8 is used as a slightly higher value.
- Gaussian standard deviation in Z (temporal) dimension (σ_t) - varied in [0.2, 1.8].
- Kernel size - in this study 5 voxels (5 x 5 x 5), 7 and 9. The number must be odd, and should not be too large as the video frame is as little in height as about 100 pixels.

The method is not expected to perform particularly well in terms of preserving the structure of an image because of its simple mechanism, which may lead to blurring the edges.

Bilateral Filtering

Bilateral filtering introduces the means to preserve the structure of the object. In addition to the tweakable Gaussian filtering parameters, the following is introduced:

- Intensity-domain standard deviation (σ_r) - to explore lower and higher sigmas varied in [0, 2], as the data is normalized.

The ability of preserving the edges is expected to have a beneficial impact on the results, and because of that improvement the method is expected to perform better than Gaussian filtering.

Anisotropic Filtering

The method diffuses the image in the direction of lower gradient. The algorithm's properties are specified by the following parameters:

- Number of iterations - varied between 3 and 20.
- Diffusion rate (γ) - varied in [0.01, 0.10]
- Edge sensitivity (κ) - determines what level of local change is treated as noise, and which is treated as an edge. A value from [0.02, 0.20].
- Temporal-spatial ratio - the ratio of temporal to spatial coordinates. Used because of the fact that the medium is a video, so it is a 3-dimensional array of intensities with an independent time dimension. The value ranged between 0.5 and 2.0.

The parameters have been chosen to span from weaker to stronger denoising, according to [13].

The method introduces a different approach to preserving edges from the bilateral filtering. It is, similarly to the mentioned approach, expected to lead to better results that Gaussian filtering.

Wavelet Filtering

The filtering method decomposes the input signal [4]. For this method, the parameters to be specified are as follows:

- Wavelet family - chosen from the standard ones: haar (square wave), db4 and sym5.

- Thresholding - soft or hard. The first means gradual thresholding, the other is applying a hard cutoff.
- Temporal-spatial ratio

Total Variation Minimization

Total variation minimization (TV minimization) reduces the number of high-frequency signals in the image. It was performed with an algorithm proposed by Chambolle [3]. There is one parameter that controls the denoising properties and the temporal-spatial ratio:

- Weight (λ) - amount of denoising, explored in the range [0,0.02], consistent with typical values used in normalized data [3].
- Temporal-spatial ratio

4.3 The Assessment Criteria

In the end, the feasibility of a denoising method is presented with a quantitative value of evaluation metrics. For the study, two metrics were proposed.

Structural Similarity Index Measure

Structural Similarity Index (SSIM) is a method to predict the quality of a filtered digital medium [18]. It puts an emphasis on perceived quality and structure of the objects on the image, as opposed to other methods (such as mean squared error, signal-to-noise ratio), which often value the absolute error instead. It requires a reference denoised image, that is, a ground truth, to calculate the measured value.

The metrics have been chosen as primary evaluation metrics due to the importance of maintaining the structure of a medical image after denoising, as only then does it allow a researcher to draw scientific conclusions from it. The measure has been used in the use case. The researchers who developed CellMincer, self-supervised denoising method developed for voltage imaging, use SSIM to evaluate their denoising method [17].

The measure can be calculated using the following formula:

$$\text{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)},$$

where μ_x and μ_y are mean intensities over a window, σ_x^2 and σ_y^2 and standard deviations, and σ_{xy} stands for covariance. The final value is calculated as the mean value of all windows. The C_1 and C_2 constants are calculated based on the pixel range.

SSIM Range	Quality Assessment
< 0.5	Poor: significant distortion
$0.5 - 0.7$	Fair: partially preserved structure
$0.7 - 1$	Good: minimal perceptible loss

Table 1: Qualitative interpretation of SSIM (Structural Similarity Index) values.

The results are compared with the values in the table 1. Values of SSIM close to 1 indicate very good to perfect structural similarity [18].

Temporal Signal-to-Noise Ratio

The evaluation metric introduced as a method to evaluate the efficiency of denoising techniques in denoising the first dataset. The dataset does not contain a ground truth, hence the need for a method that allows to approximate the amount of noise from a denoised image alone.

The signal-to-noise (SNR) ratio measure has been chosen in this research, due to its extensive use in medical imaging [10]. It is widely used for magnetic resonance imaging (MRI) scans, which require denoising, just as voltage imaging videos.

The signal-to-noise ratio is a fundamental measure that indicates the amount of noise in the video [19]. While it normally requires a ground truth to calculate the noise, there exists a way to approximate it from a video alone. The temporal signal-to-noise ratio (tSNR) is a measure that addresses the limitation of lack of ground truth for the task. It approximates SNR using the temporal dynamics of the video. It is defined as the mean signal at each pixel across time, divided by its corresponding standard deviation.

$$\text{tSNR}(x,y) = \frac{\mu_{x,y}}{\sigma_{x,y}}$$

In this study, mean tSNR for the entire video is calculated to provide a single numerical value for the comparison of the methods. Furthermore, 25th percentile of the tSNR values is evaluated, as an overview of the worst-case behavior in the problematic segments of the video.

tSNR Range	Quality Assessment
< 20	Poor: high temporal noise
$20 - 50$	Moderate: usable but noisy
$50 - 60$	Good: high quality and useful
> 60	Very Good: very low noise

Table 2: Qualitative interpretation of temporal signal-to-noise ratio (tSNR) values.

The results are compared with the values in the table 2. According to research around the SNR measure for MRI purposes, images with SNR values larger of 50-60 units are generally consider to have good quality, allowing for reliable detection of subtle signal changes [19].

4.4 Experiment Protocol

Data normalization

Before running the experiments, the intensity of all videos in normalized to ensure a consistent format, so 3 dimensional arrays of floating point numbers on scale [0, 1]. All of the videos are in grayscale, so each voxel is represented by the intensity value.

Hyperparameters selection

First, the optimal hyperparameters are selected for each of the denoising methods. The parameters are selected from the domains specified in the Denoising Approaches subsection of the Methodology. The selection is done based on the experiments performed on the validation set, by calculating the Structural Similarity Index measure. The optimal parameter values are outlined in the results section.

Evaluation

After parameter selection, the denoising methods are run on the datasets with optimal settings. The fitness of a method is later assessed by the following criteria:

- The mean SSIM over the synthetic dataset.
- Mean tSNR and 25th percentile worst-case tSNR in real neural videos.
- The qualitative evaluation of the denoised videos from the real dataset.

5 Results & Discussion

This section presents and discusses the results of the experiments, as conducted according to the protocol. It first presents the numerical results of the SSIM calculation for different settings of the denoising methods. As a result of that, the optimal parameters are selected. Furthermore, the outcomes of the experiments, as defined in the experiment protocol, are listed. Finally, the section discusses the obtained results.

5.1 Parameters Optimization

All of the methods has been run for different setting on the first 1000 frames of the optosynth__1__20__5 video.

Gaussian Filtering

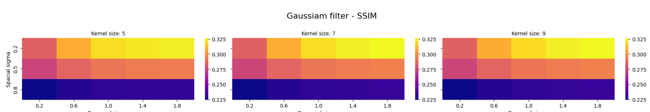

Figure 6: Gaussian Filtering - SSIM

The results of the parameters tests are presented in the Figure 6. From the figure it can be deducted that the optimal values of the parameters are the following:

- Kernel size = (9, 9, 9)
- $\sigma_s = 0.2$
- $\sigma_t = 1.8$

Bilateral Filtering

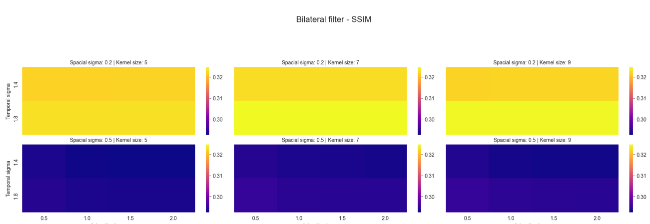

Figure 7: Bilateral Filtering - SSIM

The results of the parameters tests are presented in the Figure 7. From the figure it can be deducted that the optimal values of the parameters are the following:

- Kernel size = (9, 9, 9)
- $\sigma_s = 0.2$
- $\sigma_t = 1.8$
- $\sigma_i = 2.0$

Anisotropic Filtering

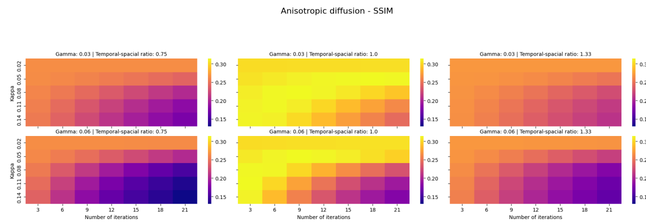

Figure 8: Anisotropic Filtering - SSIM

The results of the parameters tests are presented in the Figure 8. From the figure it can be deducted that the optimal values of the parameters are the following:

- $\kappa = 0.08$
- $\gamma = 0.03$
- 9 iterations
- Temporal-spatial ratio = 1.0

Wavelet Filtering

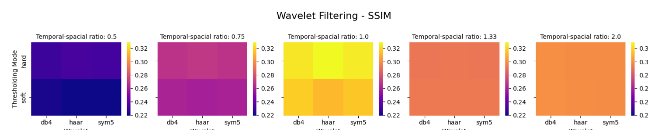

Figure 9: Wavelet Filtering - SSIM

The results of the parameters tests are presented in the Figure 9. From the figure it can be deducted that the optimal values of the parameters are the following:

- Hard thresholding
- Haar (square) wavelet
- Temporal-spatial ratio = 1.0

Total Variation Minimization

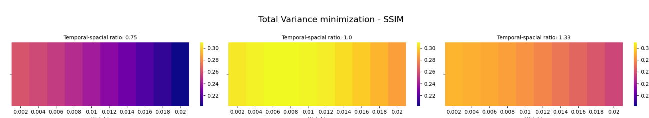

Figure 10: TV Minimization - SSIM

The results of the parameters tests are presented in the Figure 10. From the figure it can be deducted that the optimal values of the parameters are the following:

- $\lambda = 0.008$
- Temporal-spatial ratio = 1.0

5.2 Results of the Experiments

Numerical results

	SSIM	tSNR	Worst tSNR
No Filter	0.30	11.07	8.75
Gaussian Filter	0.33	23.38	23.32
Bilateral Filter	0.33	23.49	19.41
Anisotropic Filter	0.30	25.30	25.04
Wavelet Filter	0.32	14.05	13.52
TV Minimization	0.29	50.65	49.24

Table 3: Evaluation results

Table 3 shows the results of the three evaluation metrics, the mean value of the structural similarity index in the synthetic dataset, the mean temporal approximation value of the signal-to-noise ratio in the real dataset, and the worst-case 25th percentile tSNR value in the same dataset.

Example denoised frames

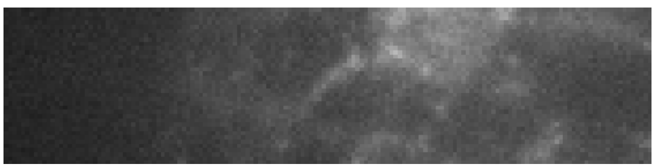

Figure 11: Noisy frame

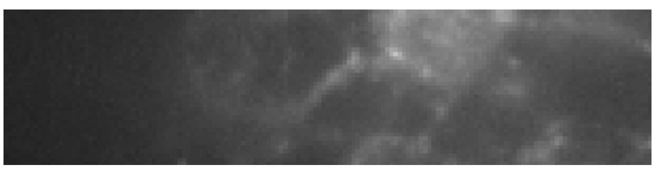

Figure 12: Gaussian Filter

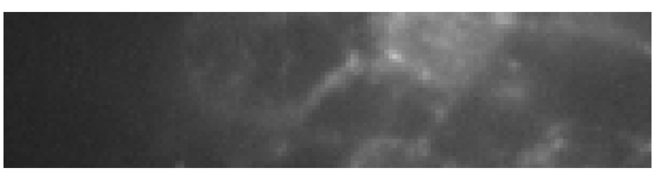

Figure 13: Bilateral Filter

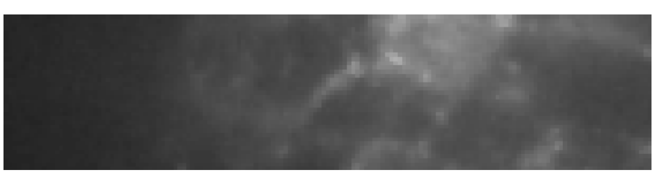

Figure 14: Anisotropic Filter

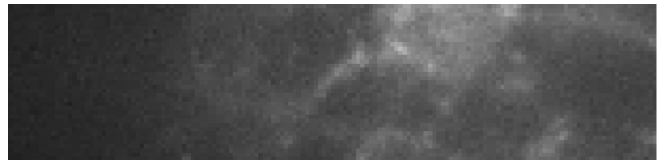

Figure 15: Wavelet Filter

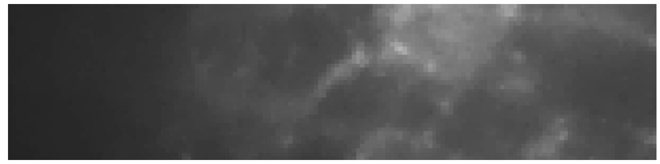

Figure 16: Total Variation Minimization

Figures 12 13 14 15 16 show the resulting frames for all the denoising methods tested. All figures show a fragment of the same (10th) frame of the *00_02* video from HPC2 dataset. The noisy frames is depicted on 11.

5.3 Discussion

For the preferred hyperparameters, the optimal parameters selected are seemingly the ones that do not interfere with the video structure, such as, for example, hard thresholding in wavelet filtering or small σ_s in the Gaussian filter, which preserves the edges in spatial dimensions. It makes sense as the parameters are tuned according to the structural similarity index. Stronger spatial sigma could blur the edges, which may be penalized by SSIM. At the same time, larger temporal deviation, blurring image over time, is preferred by the tuning algorithm, as the structure of a neuron does not change drastically across the frames.

Relatively subtle denoising is seemingly preferred by the tuning algorithm ($\kappa = 0.08$ in anisotropc filtering, $\lambda = 0.008$ in tv-minimization). Stronger spatial denoising negatively impacts the structure of the image.

Because of that, almost no improvement in SSIM is observed after denoising, which shows that denoising with the presented method inherently impacts the structure in the negative way.

As for the tSNR measures, it is visible that the noise is being removed indeed. There is no significant variance of the results, which means that the noise is being removed equally well from most of the video regions. Particularly satisfying values are noted for total variation minimization. The method is particularly good at removing high frequency noise, which is usually present in voltage imaging videos.

The denoised videos do provide a clearer overview of the action potentials happening in the neurons. The resulting videos show that anisotropic filtering and total variation minimization are particularly fit for removing the extensive noise from these scans.

In conclusion, denoising the videos does not result in a particular improvement in the structural similarity. There exists a trade-off between removing the noise and preserving the structure, particularly in Gaussian filtering, bilateral filtering, and total variation minimization. From the methods tested, total variation minimization performs particularly well in removing excessive noise. Visualizations suggest that this method, and also anisotropic filtering, are particularly suitable for making the videos more readable for researchers.

6 Responsible Research

The section serves as an overview of the ethical aspects of this study, as well as reproducibility of the methods used in the study.

6.1 Data availability

The research aims to use widely available data in order to provide maximum reproducibility for the experiments. The datasets are available online under The Creative Commons Attribution license (HPC2), and BSD 3-Clause License (CellMincer data). The open access of the data allows for the experiments described to be as reproducible as possible.

6.2 Experiment reproducibility

The research aims at maximum reproducibility of the experiments. The experimental setup of this study consists of denoising and evaluation metrics that are widely known in the field of image processing and well documented in the literature. Their implementations can be found in common image processing libraries. Similarly, evaluation metrics are described in the medical image processing literature. The formulas are precisely outlined in this paper. All improvements to the denoising methods and the specific means of their application, as well as the parameters chosen, have been well documented in this paper. The experiments do not require extensive computational power. The availability of the datasets, denoising methods, and evaluation metrics, as well as the proper description of the experiments, make the research reproducible and open.

6.3 Data interpretation

The results of the study are reliable data. The experiments have been carried out on several videos, some of which are actual neurobiological scans. The others are the media from an acclaimed synthesis tool, used in research for this type of studies. The methods tested in this study have been evaluated with reliable and established metrics used in medicine. The results and conclusions of the experiments should therefore be considered reliable and sensible.

6.4 Research limitations

The data obtained in the research do not indicate that the methods described can be used directly in healthcare applications. The experiments have not been conducted on a human brain. The images in the HPC2 dataset show only a hippocampus region of a mouse brain. If the methods described in the study are to be used for medical purposes, they should first be tested against another dataset, preferably containing scans of a human brain.

6.5 Environmental impact

The research focuses on exploring the potential of using traditional methods in the field, instead of resource-consuming deep learning methods. Adjusting these first methods to fit the use case can be potentially very beneficial to the environment. Using methods that allow one to use less computational power, such as the traditional techniques described in this paper, allows one to greatly reduce the energy consumption. Naturally, a smaller energy consumption leads to reducing carbon footprint, and as a result preserving the natural resources of the environment. In conclusion, research contributes positively to the sustainable use of natural resources and sustainable means of neuroscientific research.

6.6 Potential misuse

A potential misuse of the methods described in this research is strictly related to the ethical implications of voltage imaging. This research is not focused on the potential use of the microscopy technique, as it only outlines the improvements on the side of image processing. There may exist negative ethical implications of voltage imaging, but this research does not contribute to the technique. Instead, it merely describes and compares alternative ways to make the output videos more readable.

7 Conclusions

This section serves as a summary of the paper. It once again iterates over what has been done for the research and what the results and conclusions are. Moreover, it indicates what further steps should be taken in the research around denoising voltage imaging videos.

7.1 Research Summary

The research outlined five common traditional denoising methods: Gaussian filter, bilateral filter, anisotropic diffusion, wavelet filter, and total variation minimization. The algorithms chosen for the study are well-established standard methods of different methodological categories. All of them are used in biomedical and microscopy contexts. In addition, the paper presented the means to indicate whether a denoised voltage imaging video is useful for neuroscientific research. It introduces the structural similarity index measure and the temporal signal-to-noise ratio measure, both of which are used in research for the evaluation of denoising biomedical data.

The experiments were conducted on several videos that present actual scans of a brain. Other videos were synthesized with an acclaimed tool used in research for the development of the deep learning-based denoising method. The results obtained in the research are reliable and relevant.

The main result of the experiments is that using the traditional methods usually carries a trade-off between removing the noise and preserving the structure of an image. The methods, in particular, total variation minimization, remove noise well, but do not show a particular improvement in structural similarity. Total variation minimization, designed to remove high-frequency noise present in voltage imaging videos, and anisotropic filtering, specializing in preserving the edges, both result in particularly readable output videos. Despite a low SSIM, these two methods may nevertheless be useful for neuroscientific research.

However, for more in-depth research, some better means to denoise the videos would be needed, possibly the deep learning methods. This research enables a more extensive use of the traditional and more energy-efficient method for some of the applications in neuroscience, but deep learning methods remain relevant.

7.2 Further Research

This research shows the capabilities and limitations of some common traditional denoising methods and their performance in denoising voltage imaging videos. For the use of the presented methods in a medical context, more research should be conducted on their suitability, preferably on a human brain scan. As mentioned, deep learning methods remain relevant and for generating results more accurate in terms of structure preservation, research on the use of such methods should be conducted in future.

References

[1] Yosuke Bando, Ramdas Pillai, Atsushi Kajita, Farhan Abdul Hakeem, Yves Quemener, Hua-an Tseng, Kiryl Piatkevich, Changyang Linghu, Xue Han, and Edward S. Boyden. Real-time neuron segmentation for voltage imaging, 2023.

[2] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 2, page 60–65. IEEE.

[3] Antonin Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20(1):89–97, 2004.

[4] D.L. Donoho. De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3):613–627, May 1995.

[5] 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, September 2023.

[6] Rishikesh U. Kulkarni and Evan W. Miller. Voltage imaging: Pitfalls and potential. *Biochemistry*, 56(39):5171–5177, July 2017.

[7] Olaf Maier. Medpy - medical image processing in python. https://github.com/loli/medpy, 2020. Accessed: 2025-06-18.

[8] Peyman Milanfar. A tour of modern image filtering: New insights and methods, both practical and theoretical. *IEEE Signal Processing Magazine*, 30(1):106–128, January 2013.

[9] Sameera V. Mohd Sagheer and Sudhish N. George. A review on medical image denoising algorithms. *Biomedical Signal Processing and Control*, 61:102036, August 2020.

[10] Kevin Murphy, Jerzy Bodurka, and Peter A. Bandettini. How long to scan? the relationship between fmri temporal signal to noise ratio and necessary scan duration. *NeuroImage*, 34(2):565–574, January 2007.

[11] Yuki Nakamitsu, Masahiro Kuroda, Yudai Shimizu, Kazuhiro Kuroda, Yuuki Yoshimura, Suzuka Yoshida, Yoshihide Nakamura, Yuka Fukumura, Ryo Kamizaki, Wlla E. Al-Hammad, Masataka Oita, Yoshinori Tanabe, Kohei Sugimoto, Irfan Sugianto, Majd Barham, Nouha Tekiki, and Junichi Asaumi. Enhancing diagnostic precision: Evaluation of preprocessing filters in simple diffusion kurtosis imaging for head and neck tumors. *Journal of Clinical Medicine*, 13(6), 2024.

[12] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[13] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, July 1990.

[14] Azriel Rosenfeld. Picture processing by computer. *ACM Computing Surveys*, 1(3):147–176, September 1969.

[15] B. M. SALZBERG, H. V. DAVILA, and L. B. COHEN. Optical recording of impulses in individual neurones of an invertebrate central nervous system. *Nature*, 246(5434):508–509, December 1973.

[16] Mehdi Taassori. Enhanced wavelet-based medical image denoising with bayesian-optimized bilateral filtering. *Sensors*, 24(21):6849, October 2024.

[17] 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), December 2024.

[18] Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, April 2004.

[19] Marijke Welvaert and Yves Rosseel. On the definition of signal-to-noise ratio and contrast-to-noise ratio for fmri data. *PLoS ONE*, 8(11):e77089, November 2013.

[20] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, July 2017.

A Use of Large Language Models in this research

Through the research, LLMs have been used for multiple purposes. This section outlines the use of them with example prompts.

A.1 Paper production

To produce some parts of this paper, like LaTeX tables and figures, ChatGPT by OpenAI has been used multiple times. Example prompts that were given to the model:

- *Create a 4 rows 2 columns table in LaTeX with one row with no background, second one with red background, the other with yellow, and one with green.*
- *How do I make a LaTeX figure appear in a specific place in the text?*

Moreover, Overleaf AI assistant has been utilized to check the proper grammar of the text.

A.2 Code generation

To write the code, prompts like the following were given to ChatGPT:

- *I have a four-dimensional ndarray which I want to represent using pyplot or seaborn. Provide a piece of code for that.*

During the development process, line completion using GitHub Copilot was also used in some cases, in particular when developing the denoising algorithms.

A.3 Concepts understanding

ChatGPT has been used to help the author better understand voltage imaging, specifically in contrast to magnetic resonance imaging. It has also been used to explain some of the denoising methods, like TV minimization in detail.

A.4 Research

ChatGPT has been used in research to ensure that the base of five denoising approaches is a proper representation of all the methods available, that is, if they indeed cover multiple ways of approaching the denoising problem.