Benchmarking Self-supervised Learning for Denoising Voltage Imaging Data

- **Voltage imaging** is a powerful technique to observe fast **neuron activity**.
- Due to how fast the processes are, the resulting videos are **very noisy**.
- **Challenge:** Standard deep learning denoisers require clean and noisy image pairs for training, but such datasets are scarce for voltage imaging.
- **Solution:** Use **self-supervised models** which only need noisy data for training.

Research question

How do different self-supervised denoising models, **Noise2Void**, **AP-BSN** and **DeepVID v2**, compare in their effectiveness at denoising synthetic and real-world voltage imaging data, as measured by **PSNR**, **SSIM** and **tSNR**?

Author

Ioan Leolea

i.leolea@student.tudelft.nl

Supervisors

Nergis Tömen

Alejandro Castañeda Garcia

TU Delft

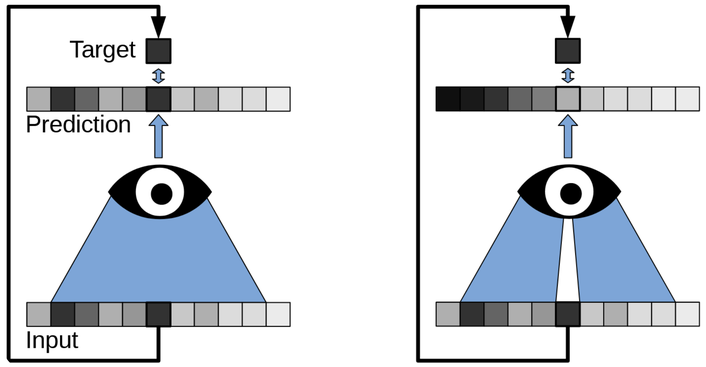
Background

Noise type

- The noise is caused by imperfections in the camera's electronics and fluctuations in the number of photons detected.
- Described by a complex Poisson-Gaussian noise distribution.

Blind-Spot Networks

- Only noisy images.
- Normal CNNs would predict the identity function.
- A BSN learns to predict a pixel's value based only on its surrounding pixels.



Methodology

Two datasets:

- One containing **synthetic images** with added Poisson-Gaussian noise.
- One containing **real images** with neurons from hippocami of mice.

Three metrics:

- Structural similarity index measure **(SSIM)**
- Peak Signal to Noise Ratio **(PSNR)**
- Temporal Signal to Noise Ratio **(tSNR)**

Three models:

- **Noise2Void:** A foundational self-supervised model used as baseline.
- **AP-BSN:** A BSN that uses subsampling to handle spatially correlated noise.
- **DeepVID v2:** A model specialized on voltage imaging data that leverages temporal information by using consecutive video frames.

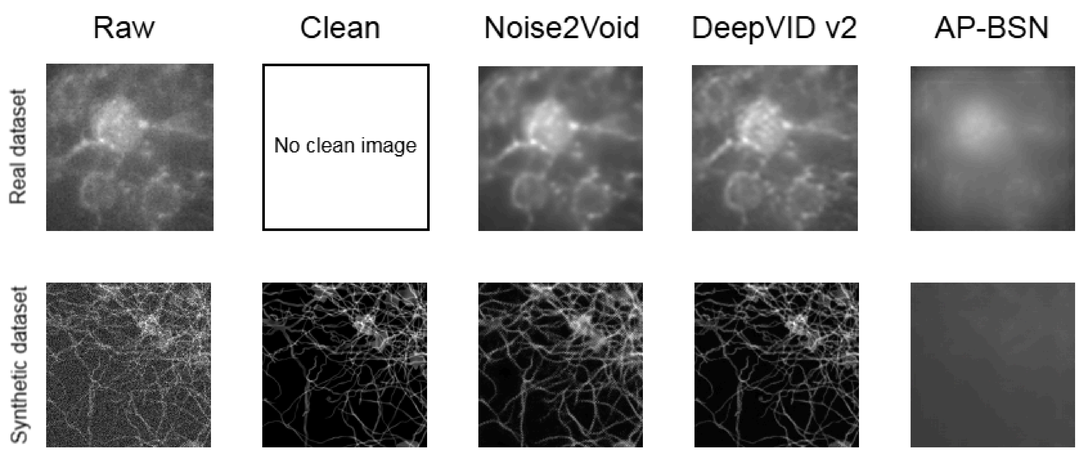
Results

Model	PSNR (dB)	SSIM
Noise2Void	26.662	0.877
AP-BSN	13.130	0.074
DeepVID v2	33.193	0.970

Synthetic dataset results

Model	tSNR (dB)
Baseline	5.36
Noise2Void	5.44
AP-BSN	6.41
DeepVID v2	5.82

Real dataset results



Limitations

- The absence of ground-truth for real data forced the use of tSNR, which proved to be an unreliable metric.
- Hyperparameters were not optimized for the models.
- A relatively small number of models was compared.

Discussion

- Self-supervised denoising is a viable strategy for voltage imaging datasets.
- **DeepVID v2** decisively outperformed the other models, because it used temporal information from videos.
- The **subsampling** technique used by AP-BSN is **ill-suited** for this specific task.
- Future work should explore **newer models** like Transformers and GANs, and develop better benchmarking datasets for real voltage imaging videos.