DENOISING OF HIGH-RESOLUTION CONFOCAL MICROSCOPY IMAGES OF DENDRITIC SPINES USING NOISE2VOID ALGORITHM

A Thesis submitted in partial fulfillment of the requirements for the degree of

Master of Technology
in
Computer Technology

Submitted by

Shyamal Kumar Surai
Registration No. : 160105 of 2021 - 2022
Examination Roll No. : M6TCT24028
Class Roll No. : 002110504034
Session : 2021 - 2024

Under the supervision of

Prof. Subhadip Basu

Department of Computer Science and Engineering
Jadavpur University
188, Raja S.C. Mallick Rd,
Kolkata - 700032, West Bengal, India

Department of Computer Science and Engineering
Faculty of Engineering and Technology
Jadavpur University, Kolkata - 700032

CERTIFICATE OF RECOMMENDATION

This is to certify that the work embodied in this thesis entitled **"Denoising of High-resolution Confocal Microscopy Images of Dendritic Spines Using Noise2Void Algorithm"** has been satisfactorily completed by **Shyamal Kumar Surai** (Registration Number 160105 of 2021 - 2022; Class Roll No. 002110504034; Examination Roll No. M6TCT24028). It is a bonafide piece of work carried out under my supervision and guidance at Jadavpur University, Kolkata, for partial fulfilment of the requirements for the awarding of the **Master of Technology** in **Computer Technology** degree of the Department of Computer Science and Engineering, Faculty of Engineering and Technology, Jadavpur University, during the academic year 2021 - 2024.

Prof. Subhadip Basu,
Department of Computer Science and Engineering,
Jadavpur University.
(Supervisor)

Forwarded By:

Prof. Nandini Mukhopadhyay,
HOD, Department of Computer Science and Engineering,
Jadavpur University.

Prof. Dipak Laha,
DEAN, Faculty of Engineering and Technology,
Jadavpur University.

Department of Computer Science and Engineering
Faculty of Engineering and Technology
Jadavpur University, Kolkata - 700032

CERTIFICATE OF APPROVAL

This is to certify that the thesis entitled **"Denoising of High-resolution Confocal Microscopy Images of Dendritic Spines Using Noise2Void Algorithm"** is a bonafide record of work carried out by **Shyamal Kumar Surai** (Registration Number 160105 of 2021 - 2022; Class Roll No. 002110504034; Examination Roll No. M6TCT24028) in partial fulfilment of the requirements for the award of the degree of **Master of Technology** in **Computer Technology** degree of the Department of Computer Science and Engineering, during the period of September 2021 to May 2024. It is understood that by this approval, the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion drawn therein, but approve the thesis only for the purpose of which it has been submitted.

Signature of Examiner 1
Date:

Signature of Examiner 2
Date:

Department of Computer Science and Engineering
Faculty of Engineering and Technology
Jadavpur University, Kolkata - 700032

DECLARATION OF ORIGINALITY AND COMPLIANCE OF ACADEMIC ETHICS

I hereby declare that the thesis entitled **"Denoising of High-resolution Confocal Microscopy Images of Dendritic Spines Using Noise2Void Algorithm"** contains literature survey and original research work by the undersigned candidate, as a part of his degree of **Master of Technology** in **Computer Technology, Jadavpur University**. All the information have been obtained and presented in accordance with academic rules and ethical conduct.
I also declare that, as required by these rules and conduct, I have fully cited and referenced all materials and results that are not original to this work.

Name: Shyamal Kumar Surai

Examination Roll No.: M6TCT24028

Registration No.: 160105 of 2021 - 2022

Thesis Title: Denoising of High-resolution Confocal Microscopy Images of Dendritic Spines Using Noise2Void Algorithm

Signature of the Candidate:

Acknowledgements

I tender my deepest gratitude and high sense of indebtedness to my guide, **Dr. Subhadip Basu**, Professor, Department of Computer Science and Engineering, Jadavpur University, Kolkata, India, for his constant support and supervision in carrying out this thesis, which has been instrumental in the materialisation of this work. His expertise and guidance was crucial in shaping both the direction and the quality of this thesis, and I am profoundly thankful for his steadfast dedication.

I express my sincere gratitude to the Head of the Department, **Dr. Nandini Mukhopadhyay**, Professor, Department of Computer Science and Engineering, Jadavpur University, Kolkata, India, for being kind enough to permit me to carry out my research at this prestigious institution.

I thank my seniors, Dr. Nirmal Das as well as my labmates, Mr. Amartya Ray, Mr. Soumyajit Gayen, Ms. Debarati Paul and Mr. Rahul for their unwavering support, validation, and understanding throughout this journey. Their companionship and camaraderie have been invaluable to me.

I must also extend a special word of thanks to my parents and family members for the encouragement and support they have rendered in preparing this thesis.

Shyamal Kumar Surai

Examination Roll No.: M6TCT24028

Registration No.: 160105 of 2021 - 2022

Department of Computer Science and Engineering,

Jadavpur University.

Abstract

The denoising is an essential pre-processing step for morphological study of dendritic spines as biomedical images contain inevitably noises due to various reason such as imperfection in microscopy, limited low light environment and slow acquisition speed (shutter speed), continuous movement of target, the consequent experimental trade-off between signal-to-noise ratio, and imaging depth. Different algorithms have been proposed in past three decades with varying denoising performances. In recent years, having outperformed all conventional methods, deep learning based models have shown a great promise in image denoising. Most of the deep learning based denoising using supervised learning methods require paired noisy and noiseless/low-noise images. But in reality, it is hard to get ground truth image. In this thesis, Noise2Void, a self-supervised model is introduced to use noisy confocal microscopy images of dendritic spine for model training and denoising and its denoising performance is assessed.

Contents

Abstract i

Table of Contents ii

List of Figures v

List of Tables ix

List of Abbreviations xi

1 Introduction 1
- 1.1 Dendritic spines . . . 2
- 1.2 Importance of Dendritic Spines Analysis . . . 2
- 1.3 Definition of Digital Image . . . 3
- 1.4 Digital Image Acquisition . . . 3
- 1.5 Noises in Digital Image Acquisition . . . 3
- 1.6 Regions of Interest (ROI) of Digital Image . . . 5
- 1.7 Confocal Microscopy . . . 5
- 1.8 Confocal Microscopy Image Acquisition . . . 7
- 1.9 Sources of Noise In Confocal Microscopy . . . 8
  - 1.9.1 Laser Noise q . . . 8
  - 1.9.2 Shot Noise(Poisson noise) . . . 8
  - 1.9.3 Multiplicative (Secondary Emission) Noise . . . 9
  - 1.9.4 Dark Noise . . . 9
- 1.10 Scope of The Thesis . . . 10
- 1.11 Organization of The Thesis . . . 10
- 1.12 Contribution of The Thesis . . . 10

2 Literature Survey 11
- 2.1 Image Denoising Era . . . 11
- 2.2 Evolution of Network For Deep Learning-based Image Denoising . . . 12
- 2.3 Evolution of Deep Learning-based Denoising . . . 15

3 Methods and Materials 42
- 3.1 Model Architecture for Noise2Void . . . 42
- 3.2 Loss Functions and optimizer . . . 46
  - 3.2.1 Mean Squared Error (MSE) . . . 46
  - 3.2.2 Mean Absolute Error (MAE) . . . 47

3.3 Network Training, Fine-Tuning, and Validation 48
3.4 Data Set 48
3.5 Metrics for quantitative and qualitative assessments of N2V denoising 48
3.5.1 Histogram 48
3.5.2 Plot Profile using Imagej Tool 49
3.5.3 SSIM 49
3.5.4 NR-IQA 50
3.6 Datasets 50

4 Results and Discussion **52**
4.1 Noise2Void 2D Trained Models 52
4.2 Experimental Results 54

5 Conclusions **64**
5.1 Scope for Future Works 64

Bibliography **66**

List of Figures

1.1 Gaussian Noise, Poisson Noise 4
1.2 Schematic of a typical confocal microscopy 7

2.1 Denoising performance on the BSD68 dataset with sigma = 25 (K-SVD[7], BM3D[5], FoE[8], LSSC[9], EPLL[10], MLP[11], CSF[12], WNNM[13], TNRD[14], DnCNN[15], IRCNN[16], NLRN[17], MVCNN[18], N3Net[19], FFDNet [20], FOCNet[21], RIDNet [22], GCDN[23], SwinIR[24], DRUNet[25]) 12
2.2 The outline of the deep learning-based image denoising methods. . . 13

3.1 The blind spot concept : Unlike a conventional network, a blind-spot network has a masked receptive field that excludes the central pixel and can learn to suppress noise by focusing on the neighboring pixels. Thus, it can generate a prediction distinct from the input even when the input and target images are identical and noisy. . . . 43

4.1 Training loss (MAE) and validation loss vs epoch number (linear scale) using Invitro Datasets 53
4.2 Training loss (MSE) and validation loss vs epoch number (linear scale) using Invitro Datasets 53
4.3 Training loss (MAE) and validation loss vs epoch number (linear scale) using Invivo Datasets 53
4.4 Training loss (MSE) and validation loss vs epoch number (linear scale) using Invivo Datasets 54
4.5 Visual denoising evaluation of unseen invitro image of dendritic spines using model trained with invitro images for both mae and mse loss functions 55
4.6 Visual denoising evaluation of unseen invitro image of dendritic spines using model trained with invivo images for both mae and mse loss functions 56
4.7 Visual denoising evaluation of unseen invivo image of dendritic spines using model trained with invitro images for both mae and mse loss functions 57
4.8 Visual denoising evaluation of unseen invivo image of dendritic spines using model trained with invivo images for both mae and mse loss functions 58
4.9 Denoising comparison Of Invitro Image of Dendritic Spines Not Used In Model Training 59

4.10 Denoising comparison of Invivo Image of Dendritic Spines Not Used In Model Training 60
4.11 Denoising comparison of Invitro Image of Dendritic Spines Having Added Noise Not Used In Model Training 61
4.12 Denoising comparison of Invitro Image of Dendritic Spines Used In Model Training 62
4.13 Denoising comparison of Invivo Image of Dendritic Spines Used In Model Training 63

List of Tables

2.1 The denoising models trained on clean/noisy image pairs using CNN image . . . 16
2.1 The denoising models trained on clean/noisy image pairs using CNN image . . . 17
2.1 The denoising models trained on clean/noisy image pairs using CNN image . . . 18
2.1 The denoising models trained on clean/noisy image pairs using CNN image . . . 19
2.2 The denoising models based on visual transformers . . . 19
2.2 The denoising models based on visual transformers . . . 20
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 21
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 22
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 23
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 24
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 25
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 26
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 27
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 28
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 29
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 30
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 31
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 32
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 33

2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 34
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 35
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 36
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 37
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 38
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 39
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 40
2.3 The denoising models trained on single/multiple image pairs (without clean) . . . 41

4.1 Details of Trained Models . . . 52

List of Abbreviations

A

AP	**A**symmetric **P**D

C

CCD	**C**harged **C**oupled **D**evice
CLSM	**C**onfocal **L**aser **S**canning **M**icroscopy
CLSM	**L**aser **S**canning **C**onfocal **M**icroscopy
CM	**C**onfocal **M**icroscopy
CNN	**C**onvolutional **N**eural **N**etwork

D

DAP	**D**eep **A**ttention **P**rior
DIP	**D**eep **I**mage **P**rior

F

FPN	**F**ixed **P**attern **N**oise

G

GAN	**G**enerative **A**dversarial **N**etworks

M

MAE	**M**ean **A**bsolute **E**rror
MIP	**M**aximum **I**ntensity **P**rojection
MM	**M**ulti-masked
MSE	**M**ean **S**quared **E**rror

N

N2I	**N**oise2Inpaint

N2K	**N**oise2Kernel
N2N	**N**oise2Noise
N2S	**N**oise2Self
N2SR	**N**oise2SR
N2V	**N**oise2Void
NAC	**N**oise-As-Clean
NAS	**N**eural **A**rchitecture **S**earch
Nei2Nei	**N**eighbour2Neighbour
P	
PI	**P**hysics **I**nformed
PMT	**P**hoto **M**ultiplier **T**ube
PN2V	**P**robabilistic **N**oise2Void
PPN2V	**P**arametric **P**robabilistic **N**oise2Void
PPRNU	**P**hoton **R**esponse **N**on**U**niformity
PSD	**P**ost **S**ynaptic **D**ensity
PSF	**P**oint **S**pecific **F**unction
R	
R2R	**R**ecorrupted-to-recorrupted
ROI	**R**egion **O**f **I**nterest
S	
SNR	**S**ignal **T**o **N**oise
SURE	**S**tein **U**nbiased **R**isk **E**stimation
T	
TV	**T**otal **V**ariation
Z	
ZS	**Z**ero **S**hot

CHAPTER 1

Introduction

Image denoising is an important pre-processing step of life science image processing as noise has been an inevitable companion of Confocal Microscopy which has become an essential toolkit in life science labs as it provides an insight into the cellular world. It is able to locate specific molecular events within cells with high spatial resolution has given insight into everything from gene expression pathways to drug interactions. Post-acquisition of images of dendritic spines images via Confocal microscopy, image denoising is an essential pre-processing step to facilitate downstream analyses such as morphological study of dendritic spines, as Confocal microscopy images contain inevitably noises due to various reason such as imperfection in microscopy, limited low light environment and slow acquisition speed (shutter speed), continuous movement of target, limited resolution (defined by the optics); uneven illumination or background; unwanted stray or out-of-focus-light reducing the image contrast; image artefacts; and, of course, noise, the consequent experimental trade-off between signal-to-noise ratio, and imaging depth.

Different algorithms have been proposed in past three decades with varying denoising performances. More recently, having outperformed all conventional methods, deep learning based models have shown a great promise. In recent years, deep learning(DL) has emerged as a very successful approach to remove this noise while retaining the useful signal. Unlike classical algorithms which use well-defined mathematical functions to remove noise, DL methods learn denoising from example data, providing a powerful content-aware approach. In this review, we first describe the different types of noise that typically corrupt fluorescence microscopy images and introduce the denoising task. In this paper, we evaluated Noise2Void[1], a novel deep learning based approach for reducing noise in images of dendritic spines (In vivo as well as In vitro) acquired by Confocal Microscopy[2] with a short exposure time.

1.1 Dendritic spines

Neuronal cells in different brain regions are covered with dendritic spines [3], which are small membranous protrusions having a length of around few microns. Dendritic spines are very plastic and their size and shape are constantly changing in response to neuronal activity. Complex machinery composed of various signaling molecules and cascades maintains the unique structure and function of dendritic spines (Yasuda, 2017; Nakahata and Yasuda, 2018). Dendritic spine shape is controlled by the actincytoskeleton. A characteristic feature of excitatory spines is a postsynaptic density (PSD), which is visible on electron microphotographs. PSD consist of densely packed ion channels, receptors, and kinases/phosphates anchored by scaffolding proteins. Learning and memory formation processes are tightly linked to remodeling or elimination of existing dendritic spines and formation of new ones, which enables modulation of information transfer efficiency between neurons (Yang et al., 2014; Segal, 2017; Chidambaram et al., 2019; Stein and Zito, 2019). For example, motor learning induces rapid growth of new dendritic spines at mice contralateral motor cortex neurons, and subsequent elimination of spines existing before training, so the overall spine density is relatively constant (Xu et al., 2009). For all these reasons, dendritic spines are believed to serve as sites for memory formation and storage, initiating memory consolidation through mechanisms of potentiation and depression of synaptic activity (Zhou et al., 2004; Bourne and Harris, 2007; Kasai et al., 2010; Bailey et al., 2015).

1.2 Importance of Dendritic Spines Analysis

Changes in dendritic spines were detected after being subject to various stimuli, including drug administration (Barrientos et al., 2018), hypoxia (Saraceno et al., 2012), environmental changes (Ashokan et al., 2018), neurodevelopmental (Nishiyama, 2019), neurodegenerative (Herms and Dorostkar, 2016) and psychiatric diseases (Penzes et al., 2011) and many others. Neurodegenerative disorders are characterized by synapse loss and dendritic spine abnormalities in the brain region associated with the disease. For example, Alzheimer's disease is known to be accompanied by dendritic spine shrinkage and elimination in the hippocampal and cortexes areas, which is proposed to start before any clinical evidence of the disease, like cognitive decline and memory dysfunction, manifests (Tackenberg et al., 2009; Boros et al., 2017). Genetic neurodegenerative disorder Huntington's disease is characterized by synapse loss in the striatal brain region, which is linked with progressive movement discoordination (Nithianantharajah and Hannan, 2013). In

contrast, autism spectrum disorders are characterized by a significant increase in spine density on various brain areas, including frontal, temporal, and parietal lobes and lateral nucleus of the amygdala (Nishiyama, 2019). Compared with healthy subjects, patients with Fragile X syndrome have elevated numbers of dendritic protrusions in the cingulate, temporal, and visual cortex with prevalence of immature one (Bagni and Zukin, 2019; Nishiyama, 2019). The balance between spine appearance, maturation, elimination, and plasticity is critical for proper brain function. Methods for analyzing and understanding the morphology of dendritic spines are critically important for many fields of neuroscience.

1.3 Definition of Digital Image

Image is defined as 2D light-intensity function indicated by *f(x,y)* where the value or amplitude of f at any space coordinates *(x,y)* depicts the intensity(brightness) of the image at that point. When x, y and the amplitude values of f are all finite, discrete quantities, we term the image as a digital image.

Digital Image comprises a fine number of elements, each with a particular value at specific location. Such elements are called picture elements, image elements and pixels. The term *Pixel* is most widely used to denote the elements of a Digital Image.

1.4 Digital Image Acquisition

Image Acquisition is the first step in any digital image processing system. This is only hardware dependent process, during which reflected light energy from the real world object being imaged is converted into electrons and cover the interior sensor chip which is sort of a 2-D array of cells. The cell is named photo-site and contain the amount of charges which is further converted to digital form using A-D Converter. Now, this digital image is used for digital image processing such as enhancement, restoration, segmentation, classifications and other manipulations.

1.5 Noises in Digital Image Acquisition

Noise is an intrinsic specificity of all forms of imaging, and can be seen in various forms in all domains of digital imagery. Noise is an unwanted artefact that appears during the image capture process and can take various forms depending on the camera model and the lighting conditions during the image acquisition. The the overall noise produced by a digital sensor is usually considered as a stationary

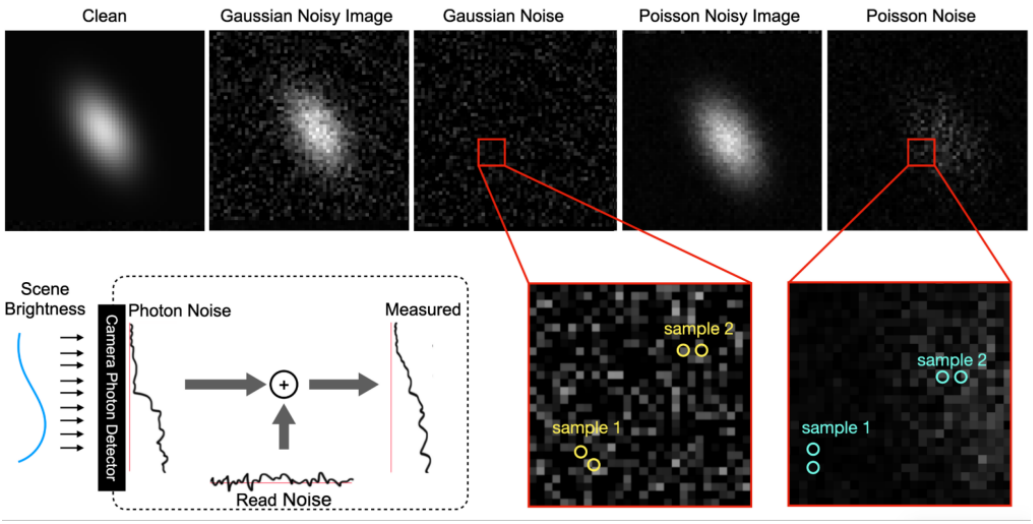


Figure 1.1: Gaussian Noise, Poisson Noise
Image Courtesy : [4]

white additive Gaussian noise.

As Additive White Gaussian Noise (AWGN) is more convenient to be modelled, most of the image denoising research focuses on AWGN removal. The AWGN denoising models assume that the noise n is the independent identically distributed Gaussian noise, which has a 0 mean and a standard variance value Although in recent years, most image denoising models are proposed for AWGN, there are some works that aim to tackle the problems in the removal of other types of noise, such as Poisson noise, and salt and pepper noise.

The main sources of noise in digital images are during the acquisition (quantity of photons collected too low, sensor temperature) or during any transmission (echoes and atmospheric distortions in wireless communication). The signal-dependent and signal-independent noise which can be approximated by Poisson distribution (Poisson noise or Shot noise) and Gaussian distribution (Gaussian noise) respectively.When photon counts are high, Poisson distribution becomes Gaussian one.

The noise being by nature a random phenomenon, however, it is modelled by a probability density which represents the intensity distribution of the noise. The Figure 1.1 depicts various noises.

The main sources of noise can be divided into two main categories :

- **Physical Noise** : linked to physics constraints like the corpuscular nature of light.
- **Hardware Noise** : linked to mechanical issues in the camera.

Physical noise notably includes dark shot noise and photon shot noise. The dark

shot noise is created by electronic fluctuations caused by an accumulation of heat-generated electrons in the sensor. It is related to thermal noise, and can be reduced by cooling down the sensor. The photon shot noise, also called Poisson noise, is the one caused by the corpuscular nature of light: as photons arrive irregularly on the photosites, two adjacent pixels supposed to have a similar value can end up with different photon counts. As the name indicates, the photons follow a Poisson distribution. Its effect decreases proportionally to exposure time.

The hardware noise includes Fixed Pattern Noise (FPN), Photon Response Non-Uniformity (PRNU) and quantification noise. PRNU and FPN are caused by imperfections in sensors. For the PRNU, the cause is mainly the inhomogeneity of silicon wafers and light variations in which individual sensor pixels convert light to electrical signals. It is most visible in pictures with a long exposure time and does not follow any particular statistical law. As for the FPN, it is caused by dark currents. Like photon shot noise, FPN tends to be reduced in long exposure images. Both of those effects increase with light intensity. While the FPN can be removed by subtracting the dark frame, the PRNU is non-linear and as such is very hard to remove.Quantification noise is caused by the analogic-numeric converter. It is hard to quantify, because the process is non-linear, though there are some accepted models.

1.6 Regions of Interest (ROI) of Digital Image

An ROI is a subset of an image appropriate for the intended analysis and identified manually by experts. ROI usually means the meaningful and important regions in the images. The use of ROI can avoid the processing of irreverent image points and accelerate the processing. Extraction of regions of interest from images is an important and unsolved topic in the image processing area, especially in biomedical image processing area.

1.7 Confocal Microscopy

Confocal microscopy(CM)[2], most frequently confocal laser scanning microscopy (CLSM) or laser scanning confocal microscopy (LSCM), is an optical imaging technique for increasing optical resolution and contrast of a micrograph by means of using a spatial pinhole to block out-of-focus light in image formation. CM is a specialized form of standard fluorescence microscopy (also called widefield fluorescence microscopy) that uses fluorescence optics to generate high-resolution images of material stained with fluorescent probes. Confocal microscopy offers several advantages over conventional optical microscopy, including shallow depth of field,

elimination of out-of-focus glare, and the ability to collect serial optical sections from thick specimens. In the biomedical sciences, a major application of confocal microscopy involves imaging either fixed or living cells and tissues that have usually been labeled with one or more fluorescent probes. In vivo application of CM is limited by the depth of single-photon tissue penetration; however, CM was successfully used for imaging of the human cornea and mouse cerebral cortex.

Confocal microscopy differs from conventional widefield fluorescence microscopy in that the optical path is designed to place in front of the image detector (photomultiplier tube or camera) an aperture (opening) at a point where the image is focused in conjunction with the focal plane of the image. Instead of illuminating the whole sample at once, laser light is focused onto a defined spot at a specific depth within the sample. This leads to the emission of fluorescent light at exactly this point. A pinhole inside the optical pathway cuts off signals that are out of focus, thus allowing only the fluorescence signals from the illuminated spot to enter the light detector. By scanning the specimen in a raster pattern, images of one single optical plane are created. 3D objects can be visualized by scanning several optical planes and stacking them using a suitable microscopy deconvolution software (z-stack). It is also possible to analyze multi-color immunofluorescence stainings using state-of-the-art. confocal microscopes that include several lasers and emission/excitation filters.

CM increases effective spatial resolution by eliminating out-of-focus light, using a point illumination source paired with a pinhole detector aperture. Only the fluorescence from the focal plane reaches the detector, usually a charge-coupled device (CCD) camera. Effective axial resolution is greater than in conventional microscopes, but longer exposures are often required due to decreased signal intensity. The followings are the main features and benefits of confocal microscopy :

- It is capable to produce high-contrast and high-resolution images by eliminating unnecessary scattered light from outside the focal plane and suppressing stray light, confocal microscopy improves lateral resolution by 30%.
- Maximum Maximum Intensity Projection (MIP) images can be obtained by using confocal microscopy because of its high resolution and deep depth of focus.
- It enables Non-invasive Observation without touching the sample, without damaging soft sample, without any pre-treatment such as addition of chemicals or processing of cells, which is necessary when using electron microscopes.

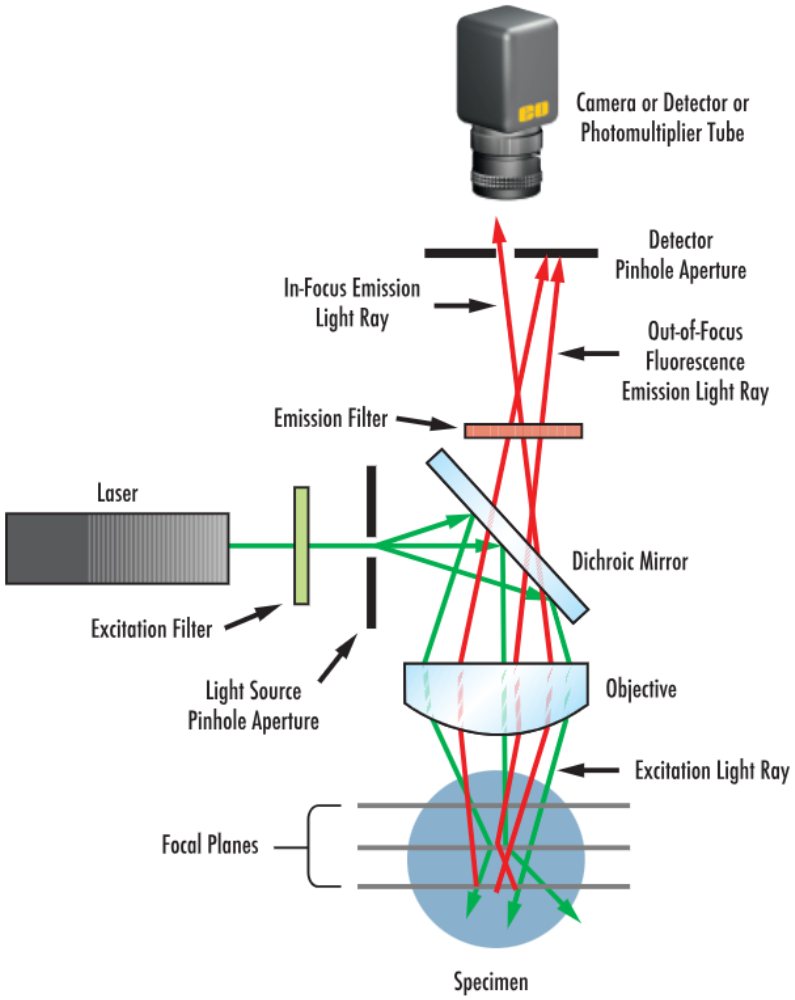


Figure 1.2: Schematic of a typical confocal microscopy
Image Source : [2]

- Three-dimensional (3D) imaging can be also performed using confocal microscopy.

1.8 Confocal Microscopy Image Acquisition

In fluorescence CLSM, sub-cellular structures are labeled with fluorescent dyes called fluorochromes. The structure to be imaged can thus be considered as a spatial fluorochrome distribution. Figure 1.2 depicts the typical layout of a confocal microscope.

A 3D image of such a specimen is acquired as follows: A laser is focused by the optics onto a point within the specimen excitation beam path). The illumination light is directed (scanned) onto the specimen dot by dot through a mirror on galvon-motor-driven scanner. The emitted light from specimen is collected and de-scanned dot by dot then passes through detecting pinhole and reaches photo

multiplier tube (PMT). So, both illumination and image detection depend on the scanning process, i.e. the image is generated through scanning. An excited fluorochrome emits fluorescence which is having longer wavelength than excitation light (Stroke's law), is collected and de-scanned do by do then passes through detecting pinhole and reaches PMT.

The emission photon flux from a single fluorescent molecule depends on the excitation photon flux, the quantum yield of the fluorochrome and its molecular cross-section. It gives the number of photons emitted per unit time. However, not only a single fluorescent molecule contributes to the signal formation of one pixel, but rather all fluorescent molecules contained in the sensing volume form the input to the detection beam path, the radiant flux L. The number of fluorescent molecules in the sensing volume is determined by the fluorochrome concentration. The sensing volume is defined by the effective support of the PSF of the excitation beam path: It is the small volume within the specimen, that is illuminated by the laser during the pixel time.

1.9 Sources of Noise In Confocal Microscopy

The main factors of noise in confocal LSM system are detector noise (dark noise, secondary emission noise), laser noise and shot noise of light. These sources of noise are of a statistical nature. Periodic noise occurs rarely ever, and if it doe, it tends to correlate with defective devices or mechanical vibration in the setup, therefore it is not considered here. The sources of noise effective confocal LSM occurs everywhere in the signal train - from the laser unit right up to A/D conversion. Essentially, four sources of noise can be distinguished as follows :

1.9.1 Laser Noise q

Laser noise is caused by random fluctuations in the filling of excited states in the laser medium. It is proportional to the signal amplitude N and therefore significant where a great number of photons (N $<$ 10000) are detected.

1.9.2 Shot Noise(Poisson noise)

This is caused by the quantum nature of light. Photons with the energy **h.v** hit the sensor at randomly distributed time intervals. The effective random distribution is known as Poisson distribution. Hence,

$$\Delta N_{Poisson} = \sqrt{N} \quad SNR_{Poisson} = N/\sqrt{N}$$

where N = number of photons detected per pixel time(= photon electrons = electrons release from the PMT cathode by incident photons). With low photon electron numbers ($N < 1000$), the number N of photons incident on the sensor can only be determined with a certainty of $\pm\sqrt{N}$
N can be computed as :

$$N = \frac{Photonenergy}{QE_\lambda . Pixeltime}$$

where QE_λ = quantum yield of the sensor at wavelength λ; 1 photon = h. c/λ; c = light velocity; h = Planck's constant

1.9.3 Multiplicative (Secondary Emission) Noise

Multiplicative noise is caused by the random variation of photo electron multiplication at the dynodes of PMT. The amplitude of multiplicative noise is a factor between 1.1 and 1.25, depending on the dynode system and the high voltage applied (gain). In general, the secondary emission noise is the lower, the higher the PMT voltage is, because a higher voltage across the dynodes improved the collecting the efficiency and reduces the statistical behaviour of multiplication.

1.9.4 Dark Noise

Dark noise is due to the generation of thermal dark electrons N_d irrespective of whether the sensor is irradiated. N_d statistically fluctuates about $\sqrt{N_d}$. Dark noise is specified for a PMT voltage of 1000V; with lower voltages it progressively loses significance. Dark noise can reduced by cooling the sensor. However, the reduction is significant only if $N \leq N_d$, e.g. in object-free areas of a fluorescence specimen. In addition, the dark noise must be the dominating noise source in order that cooling effects a signal improvement, in most applications, the will not be the case.

Additional source of noise to be considered are amplifier noise in sensor diodes and readout noise in CCD sensors. In the present context, these are left out of considerations.

The mean square deviation $\Delta N_{Poisson}$ from the average of $N + N_d$ of the photo-electrons and dark electrons registered,

$$\Delta N_{Poisson} = se.\sqrt{(N + N_d)(1 + q^2)}$$

so that the total signal-to-noise ratio can be given as

$$SNR = \sqrt{\frac{N^2}{(se)^2(N+N_d)(1+q^2)}}$$

where

N	number of photo-electrons per pixel time (sampling time)
se	multiplication noise factor of secondary emission,
q	peak-to-peak noise actor of the laser,
N_d	number of dark electrons in the pixel or sampling time

1.10 Scope of The Thesis

This thesis makes significant contributions to denoising confocal microscopy images of dendritic spines using Noise2Void[1], a novel self-supervised image denoising model.

1.11 Organization of The Thesis

The remainder of this thesis is organised as follows :

- The Chapter 2 presents a comprehensive literature study about image denoising approaches starting from traditional era to AI Revolution Era.
- The Chapter 3 illuminates methods and materials are used for denoising of invivo and invitro images using Noise2Void.
- The Chapter 4 elucidates about results and analysis of performance of denoising capability of Noise2Void for confocal images of dendritic spines.
- Chapter 5 concludes this thesis with a future prospects of denoising of confocal images of dendritic spines using Noise2Void.

1.12 Contribution of The Thesis

In this thesis, Noise2Void self-supervised image denoising model is introduced for denoising noisy confocal images of dendritic spines (invivo and invitro specifically) without possession of ground truth or clean image and its denoising performance is experimented, analysed and evaluated.

CHAPTER 2

Literature Survey

In this chapter, we have an aim to completely review and summarize the deep learning technologies for image denoising proposed in recent years. Moreover, Image denoising era starting from classical era to AI revolution era of denoising are described.

2.1 Image Denoising Era

The progress in image denoising with accumulated knowledge in image processing and design of image denoising methodologies can be classified into two eras:

- **Classical Era of Denoising :** The classical era of denoising started in 70's and ended in the past decade. The years 1980 – 2010 have seen consistently improving denoising algorithms, many of which relying on the Bayesian point of view. This progress has been geared by an evolution of image priors that form the backbone of the overall progress in image processing and eventually led to many well-performing denoising algorithms. The classical era, started with the early L2-based regularization, proceeding to robust statistics, moving to the introduction of wavelets, and the later deployment of partial differential equations to imaging tasks, and this continued all the way to sparse modelling, patch-based methods, and low-rank structure assumptions. The famous benchmark denoising methods in classical era are as follows:
  - Non-Local Self Similarity [4]
  - Non-Local means (NLM) [4]
  - Block-matching and 3D filtering (BM3D) [5]
  - Weighted nuclear norm minimization (WNNM) [6]

- **AI Revolution Era of Denoising :** This era started around 2012 and is very much vivid till this day. The past decade has brought a paradigm shift to the general topic of data processing due to the emergence of the Artificial Intelligence (AI) revolution. In recent years. the deep learning-based image denoising approach started gaining popularity, as it is able to overcome some drawbacks existing in the classical/conventional denoising methods, such as sophisticated optimization procedures, manually set parameters, and certain models working effectively only for certain tasks. Interestingly, few popular denoising algorithms continues to serve as the foundations for building deep learning based denoising, even if utilized implicitly. So, conventional denoising algorithm explored in classical era are not dead even in deep learning denoising era.

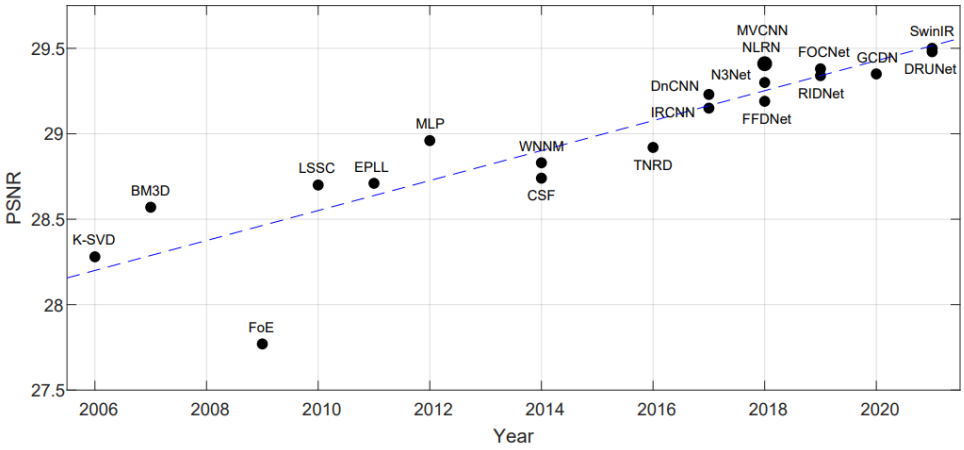


Figure 2.1: Denoising performance on the BSD68 dataset with sigma = 25 (K-SVD[7], BM3D[5], FoE[8], LSSC[9], EPLL[10], MLP[11], CSF[12], WNNM[13], TNRD[14], DnCNN[15], IRCNN[16], NLRN[17], MVCNN[18], N3Net[19], FFDNet [20], FOCNet[21], RIDNet [22], GCDN[23], SwinIR[24], DRUNet[25])
Figure Courtesy : [4]

2.2 Evolution of Network For Deep Learning-based Image Denoising

Improving image denoising capabilities came to a natural new front, where the aim is to perform better in terms of Peak-Signal-to-Noise-Ratio an agreed-upon corpus of test images. This is manifested by an evolution of architectures that started with simple feed-forward Convolutional Neural Networks(CNN) [15], proceeded to more advance structure such as U-Net and all the way to recently

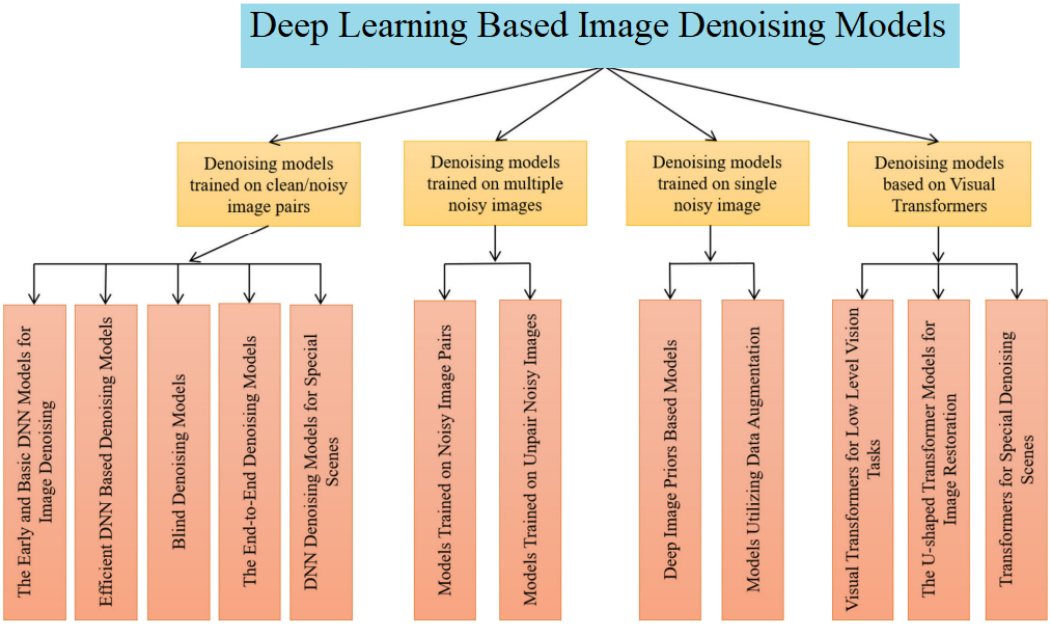


Figure 2.2: The outline of the deep learning-based image denoising methods. Figure Courtesy : [4]

introduced Transformers, more specifically Vision Transformer(ViT) [26] and Diffusion model has introduced DDPMs[27].

The Figure 2.1 illustrates the trend by presenting a graph that shows the progress in denoising PSNR on the well-known BSD68 dataset.

The deep learning-based image denoising methods are roughly divided into four categories on basis of training or learning process as follows :

- Trained on clean/noisy image pairs
- Trained multiple noisy images
- Learned from a single noisy image
- Learned from Transformer based image restoration methods.

The categories of deep learning-based image denoising methods are shown in Figure 2.2.

The network of deep learning based denoising are predominantly categorised as follows :

- **Convolutional Neural Networks (CNN) [Since 2012] :** Since deep learning popularity began in computer vision field in 2012, the Convolutional Neural Networks(CNN) have been default backbone architecture. It was the rule for any image denoising researcher to customise a model based on CNN to increase the efficiency of the state of the artwork on his task. Burger

et al. showed that a well-trained multi-layer perceptron (MLP)[11] network over a massive collection of noisy and clean patches could outperform the well-known BM3D[5] method. TNRD[14] by Chen et al. which exploited advances in partial differential equations for image restoration. TRND designed a flexible denoising framework in which each stage was modelled by a convolutional layer with large trainable kernels optimized over a large dataset. Notably, DnCNN by Zhang[12] introduced the earliest attempt in this avenue and exhibited superior performance with simpler architectures. Then DnCNN has been successfully employed in many model-based denoising algorithms serving as an implicit natural image prior. However, CNNs have several limitations regarding their ability to learn mapping from degraded images to reach their original image. In addition, the quality of the image generates by CNNs is often far from the original images.In contrast, Generative Adversarial Networks (GANs) and Diffusion model address challenges of CNN.

- **Generative Adversarial Network (GANs) [Since 2014] :** These are a class of deep learning models that involve training two neural networks, a generate and a discriminator to work together to generate new and synthetic data like the real data. Despite the success of GANs in deep learning, their applications in image denoising methods have often found to produce the sub-optimal results due to issues such as pattern collapse, excessive smoothing, artifacts, and training instability.

- **Vision Transformer(ViT) [Since 2020] :** Vision Transformer model is a deep natural architecture based on Transformer architectures has been remarkably successful in image restoration. ViT in the new architectures for image denoising is becoming a rule. This is due to some advantages compared to CNN, such as better efficiency, especially when more data are fed to network, robustness in feature extraction, and a better feature learning approach that sees better the variances and characteristics of the input. Nevertheless, some drawback exist, such as the need for more data to show benefits of ViT over CNN, the increased computation cost due to complexity of the self-attention block, a more challenging training process and lack of interpretability. Liang et al. proposed the SwinIR [24] based on the Swin Transformer[26], the SwinIR has achieved state-of-the-art performance in many image recovery tasks, including image denoising.

- **Denoising Diffusion model [Since 2020] :** The ideas of denoising diffusion model borrows concepts from the probabilistic methods such as Markov

Chains. This model is deep generative model which has two step processes: forward diffusion process to add noise (Gaussian noise) to the available training data and then reverse diffusion process/reconstruction process undoes the noise by learning the conditional probability densities using a neural network model. Diffusion probabilistic models [94], first introduced by Sohl-Dickstein et al., are a class of generative models that aim to match a given data distribution by learning to reverse a gradual, multi-step noising process. Ho et al. proposes a specific parameterization of the generative models that simplifies the training process and establishes the equivalence between denoising diffusion probabilistic models (DDPMs)[27] and score-based generative models, which learn a gradient of the log-density of the data distribution using denoising score matching[28]. The DDPMs are generative models mainly used remove noise from visual or audio data. In contrast to transformers, which require large datasets for effective training, diffusion models require much less data for effective training. This make diffusion models ideal for situations in which data may be sparse or limited. It can be trained quickly. These models are ore computationally efficient than transformers. However, high computational cost and generating images with accurate global feature are limitation of diffusion model.

2.3 Evolution of Deep Learning-based Denoising

Traditionally, from training perspective, the most of deep learning-based image denoising approaches are supervised methods such as Noise2True that require training pairs of noisy images and their respective clean target images (ground truth). However, getting clean counterpart of images acquired from imaging devices is not trivial on most of the applications including microscopy, fluorescence, medical imaging, and even normal photography. Also, obtaining clean images is not possible at all if the image details are especially volatile. Due to this reason, Lethinen et al offers an elegant solution Noise2Noise (N2N)[29] for non-availability of ground truth. It does not need noise free images (ground truth) but can be trained using pairs of noisy images, both being pictures of the same scene. That is two images s + n1, s + n2 containing the same true signals but infected by two different, independently drawn noises n1 and n2. The idea is to train a CNN to use one of the noisy images to try to predict the noisy pixel values of other images. It may sound counter intuitive to predict noisy values when the goal is to denoise an image. The basis of the method is that since the noise has an expected value of zero, the expected values being minimized during training are same regardless if the target image is clean or noisy. Because of this, the prediction that

minimizes the expected loss ends up being a denoised version of the image. Intuitively, we can view this method as the pixels being "correct on average" even if they are corrupted by noise, so that they can still be used for learning. Despite of the success of Noise2Noise, these methods require pair of noisy images. The acquisition of such pairs with (quasi) constant s is possible for static scenes (quasi). Note that Noise2Noise is basically still a supervised method, just with noisy supervision. In recent years, self-supervised methods such as Noise2Void[1], Noise2Self[30] for image denoising is increasingly popular because it does not require any prior knowledge for more general cases where only the individual noisy images are available without any extra information. Self-supervised methods are such as, the Noise2Void[1], Noise2Self [30] and Noise2Same[31]. Most of the existing self-supervised methods are based on convolutional neural networks(CNN), which are restricted by the locality of the receptive field and would cause colour shifts or textures loss. Transformer based denoising models are proposed such as SwinIR[24], SwinIA[32]. Moreover, the recent trend in denoising using denoising diffusion models based on generative model DDPMs[27] are also gaining more attention to overcome limitation of CNN and Transformer.

The literature survey of deep learning-based image denoising models covering three areas

- The denoising models trained on clean/noisy image pairs using CNN is shown in Table 2.1.
- The denoising models based on visual transformers is shown in Table 2.2.
- The denoising models trained on single/multiple image pairs (without clean) is shown in Table 2.3.

Table 2.1: The denoising models trained on clean/noisy image pairs using CNN image

Model Name	**Model Details**	**Area of Applications**
DnCNN[15]	CNN with residual learning and BN for image denoising	Gaussian image denoising, Image super-resolution
FFDNet[20]	CNN with tunable noise level for image denoising	Gaussian image denoising, Real noisy image denoising
DRUNet[25]	CNN with U-Net and ResNet for image denoising	Gaussian image denoising, Image deblurring, Image super-resolution, image demosaicing

Table 2.1: The denoising models trained on clean/noisy image pairs using CNN image

Model Name	**Model Details**	**Area of Applications**
BUIFD[33]	CNN with Bayesian estimation theory for image blind denoising	Gaussian image denoising
UNet,UNLNet [34]	CNN with local and non-local operators for image denoising	Gaussian image denoising, Real noisy image denoising
NLNet[17]	CNN with non-local operator for image denoising	Gaussian image denoising
SGN[35]	CNN with a top-down self-guidance strategy for image denoising	Gaussian image denoising, Low-light image denoising
DIPNet[36]	CNN with residual blocks for image blind denoising	Gaussian image denoising, Real noisy image denoising
HiNAS[37]	CNN with dilated convolution and skip connection for image denoising	Gaussian image denoising
DDFN[38]	CNN with deep boosting framework for image denoising	Gaussian image denoising, JPEG image deblocking, Real noisy image denoising
FOCNet[21]	CNN with the discretisation of a fractional-order differential equation for image denoising	Gaussian image denoising
CSConv[39]	CNN with class specific convolution for image denoising	Gaussian image denoising
DCDicL[40]	CNN with dictionary learning for image denoising	Gaussian image denoising
CBDNet[41]	CNN and in-camera processing pipeline for real noisy image blind denoising	Real noisy image denoising
SDNet[42]	CNN with lifted residual blocks for image blind denoising	Gaussian image denoising, Real noisy image denoising
RIDNet[22]	CNN with skip connection and feature attention for image denoising	Gaussian image denoising, Real noisy image denoising

Table 2.1: The denoising models trained on clean/noisy image pairs using CNN image

Model Name	**Model Details**	**Area of Applications**
D-BSN[43]	CNN with self-supervised learning and knowledge distillation for image denoising	Gaussian image denoising, Real noisy image denoising
DUBD[44]	CNN with conditional affine transform blocks and residual blocks for image blind denoising	Gaussian image denoising, Real noisy image denoising
NBNet[45]	U-Net with subspace attention for image denoising	Gaussian image denoising, Real noisy image denoising
SADNet[46]	CNN with residual spatial-adaptive block for image denoising	Gaussian image denoising, Real noisy image denoising
InvDN[47]	CNN with wavelet transformation and invertible block for image denoising	Real noisy image denoising
P3AN[48]	CNN with pseudo 3D auto-correlation attention blocks for image denoising	Gaussian image denoising, Real noisy image denoising
DeamNet[49]	CNN with adaptive consistency prior constraint and dual element-wise attention mechanism module for image denoising	Gaussian image denoising, Real noisy image denoising
HiNet[50]	CNN with half instance normalisation block and residual block for image restoration	Real noisy image denoising, Image deblurring, Image deraining
DeepFnF[51]	Encoder-decoder with pooled skip connection for image denoising	Low-light image denoising
HORUS[52]	U-Net and a physical noise model for image denoising	Low-light image denoising
Hwnetf[53]	CNN with pseudo-3d block for image denoising	Hyperspectral image denoising

Table 2.1: The denoising models trained on clean/noisy image pairs using CNN image

Model Name	**Model Details**	**Area of Applications**
MAXIM[54]	Encoder-decoder with multi-axis gated MLP block, residual channel attention block and Cross Gating Block for image processing	Real noisy image denoising, Image deraining, Image debluring, Image dehazing, Image detouching
AirNet[55]	NN with contrastive-based degradation encoder, degradation-guided groups and degradation-guided module for image restoration	Gaussian image denoising, Image deraining, Image dehazing
FADNet[56]	Encoder–decoder with Fourier transform and plane dynamic filter module for image denoising	Real noisy image denoising

Table 2.2: The denoising models based on visual transformers

Model Name	**Model Details**	**Area of Applications**
Uformer[57]	Encoder–decoder with transformer blocks for image restoration	Real noisy image denoising, image deraining, image debluring, Moire pattern removal
IPT[58]	Encoder–decoder transformer for image restoration	Gaussian image denoising, image deraining, image super-resolution
SwinIR[59]	NN with residual swin transformer blocks for image restoration	Gaussian image denoising, image super-resolution, JPEG compression artifact reduction
Restormer[60]	NN with multi-Dconv head transposed attention, Gated-Dconv feed-forward network and transformer blocks for image restoration	Gaussian image denoising, real noisy image denoising, image deraining, image debluring

Table 2.2: The denoising models based on visual transformers

Model Name	**Model Details**	**Area of Applications**
U2-Former[61]	Encoder–decoder with transformer blocks for image restoration	Image reflection removal, image deraining, image dehazing
Rformer[62]	Transformer with generator and discriminator for fundus image restoration	Restored fundus images
TransWeather[63]	Transformer encoder–decoder for image restoration	image restoration of weather degradation
SiamTrans[64]	CNN encoder–decoder and transformer with temporal and spatial attention modules for image restoration	Image deraining, image demoireing, image desnowing
VRT[65]	Transformer with temporal mutual self-attention and parallel warping for video restoration	Video super-resolution, video debluring, Gaussian video denoising, video frame interpolation

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
DIP[66]	Supervised, DnCNN	Nov 2017	Image blind denoising, Image inpainting, Image super-resolution, Flash-no-flash image reconstruction	This is novel model, but takes much time to adjust results, there is no superior strategy to determine when to stop. The weaknesses are Over parameterized, Overfit to noise if run for too many iterations, not suitable for non-zero mean noise, performance is sensitive to moment of stopping its network training, so premature stopping leading to over-smooth results with degraded image.
N2N[67]	Supervised,U-Net	Mar 2018	Gaussian image denoising, Fluorescence microscopy image denoising	It requires pair of noisy images for same clean image, which is hard to acquire for biomedical image. It utilizes a mean squared distance error (L2 loss)

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
FC-AIDE[68]	Supervised, DnCNN-B & DnCNN-S	Jul 2019	Gaussian image denoising, Medical image denoising	It uses heavier network as having 820 parameters. LIDIA Outperform N-AIDE, Possible future extending towards to other general image restoration problems, beyond denoising and devising a full blind denoising.
N-AIDE[69]	Supervised,CNN	Sep 2018	Gaussian image denoising, Medical image denoising LIDIA	FC-AIDE Outperform N-AIDE
DIP-TV[70]	Supervised, DnCNN	2019	Image blind denoising, Image inpainting, Image super-resolution, Flash-no-flash image reconstruction, Combining DIP with TV denoising	It combines DIP with TV (total variations) denoising and outperforms the DIP and TV

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
N2V[1]	Self-Supervised, Residual UNet using MaxPool and blind spot strategy	Apr 2019	Gaussian image denoising, Fluorescence microscopy image denoising	It requires only single noisy image so benefiting over N2N. Its Blind-spots replacement scheme leads to visible checkerboard artifacts. Its high frequency suppression techniques to counteract strongly. It cannot suppress correlated or structured noises. It does not utilize self-similarities / Symmetries in an image as NLM method. Its training scheme ignores the center pixel masking scheme as results few pixels contribute to loss function. It utilizes a mean squared distance error (L2 loss)
GAN2GAN[71]	Supervised, GAN	May 2019	Gaussian image denoising, Mixture and correlated noise removal, Microscopy image denoising	It requires pair of noisy images for same clean image, which is hard to acquire for biomedical image.

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
Noise2Self[72]	Self-Supervised, DnCNN	Jan 2019	Natural image denoising, Chinese character image denoising, Fluorescence microscopy image denoising	Its training scheme ignores the center pixel masking scheme as results few pixels contribute to loss function. It cannot suppress correlated or structured noises.
Laine et al.[73]	Self-Supervised, U-Net	Jan 2019	Gaussian image denoising, Poisson image denoising, Impulse image denoising	It requires pair of noisy images for same clean image, which is hard to acquire for biomedical image.
eSURE[74]	Supervised, DnCNN, CD-nCNN	Feb 2019	Gaussian image denoising, SURE-based image denoising, Monte-Carlo-based image denoising	It requires pair of corelated noisy images for same clean image, which is hard to acquire for biomedical image. Useful for imperfect ground truth images with mild noise. It outperforms SURE, Noise2Noise.

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
Deep-HS-prior[75]	CNN	Mar 2019	Hyperspectral image denoising, Image inpainting, Image super-resolution	Due to limited availability of hyperspectral training samples and the high dimensionality of spectral bands, it is difficult to fully exploit spatial information and correlation among the spectra of hyperspectral data.
NAC[76]	Self-Supervised, CNN	Jun 2019	Gaussian image denoising, Real noisy image denoising	It is computationally efficient, improves image quality smoothing, contrast boundary sharpness.
PN2V[77]	Self-Supervised, U-Net with blind-spot strategy	Feb 2020	Gaussian Mixture Models based noise denoising, Real world microscopy image denoising	It is fully probabilistic approach extending self-supervised CARE training and N2V via mask convolution. It transfers the masking procedure to a feature level for larger receptive fields, which reduces valuable information loss. It needs an additional noise model for which calibration data needs to be acquired.

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
PPN2V[78]	U-Net	Mar 2020	Parametric Gaussian Mixture Models based noise denoising, Real world microscopy image denoising	It improvises PN2V by replacing histogram-based noise models by parametric noise models based on Gaussian Mixture Models (GMM) and bootstrapping a suitable noise model even in absence of calibration data.
NAS-DIP[79]	Self-Supervised, U-Net	Aug 2020	Image blind denoising, Image inpainting, image super-resolution, flash-no flash image reconstruction, Content Aware Filling, content up scaling, unpaired image-to-image translation.	It combines NAS with DIP. The drawback of reinforcement learning based NASS is that they need additional training to evaluate design choices. It impose a significant time overhead on the search process. Another problem is that NAS approaches rely on discrete choices performed and exploited by controller. This can be omitted by using gradient based search to build a flexible hierarchical search space.

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
Nr2N[80]	Self-Supervised, U-Net	Jun 2020	Gaussian image denoising, multiplicative Bernoulli noise removal	The main assumption of this is that noise from the same noise statistic as the original noisy image should be injected into the noisy images so that the neural networks learn to reduce the noise from the noisy input image.
Noise2Same[81]	Self-Supervised,U-Net	Oct 2020	Gaussian image denoising, Chinese character image denoising, microscopy image denoising	It makes two forward passes through (one with a random mask, one without) and regularizes the training with invariance loss in masked pixel locations.

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
Self2Self[82]	Self-Supervised, Set of NNs whose weights are scaled by Bernoulli distributions (with encoder blocks and decoder blocks)	Jun 2020	Gaussian image denoising, salt-and-pepper image denoising, Real noisy image denoising, Image inpainting.	It proposes novel framework with dropout using Bernoulli-sampled instances of single input noisy image. It does not utilize prior information about on noise distribution or level. Its drawbacks are as follows : Long denoising time, poor performance for low noise level, heavily on ensembling of multiple networks, prediction og different brightness for salt-and-paper noise leads to appearing brightness shifting artifact always in the case of highly corrupted non-zero mean noise.
Noise2Stack[83]	Either Supervised or self-supervised, U-Net	Nov 2020	Gaussian image denoising, fluorescence microscopy image denoising (3D), low dose computed tomography	It extends Noise2Noise for volumetric data.

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
Noise2Sim[84]	Self-Supervised, U-Net	Nov 2020	Gaussian image denoising, Fluorescence microscopy image denoising	It is first similarity-based unsupervised deep approach using noisy images only suppressing not only independent but also correlated noises (drawback of N2V, N2S). It finds similarity-based training samples satisfying the ZCN (Zero-mean conditional noise) and ZCD (zero-mean conditional discrepancy) conditions from noisy data without any labels. A shift towards low frequency components for Noise2Sim arising from blurring of tissue boundaries and residual image transfer induced by masking of dissimilar regions in the loss function. It recovers intrinsic features from noisy low-dose CT and photon-counting CT images.

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
N2I[85]	Self-Supervised, Shallow and deeper U-Net	Nov 2020	Both Structured and Unstructured image denoising, Real noisy image denoising	It regularises image inpainting framework. It can deal with structured noise in biomedical images. It first recast the denoising problem with holdout self-supervision as an iterative regularized inpainting problem consisting of data fidelity and regularizer terms, and then unroll the iterative algorithm for fixed number of iterations. The training of this network was performed end-to-end by partitioning the noisy image pixels into two disjoint sets, similar to the purely data-driven Noise2Self, where one set was utilized in the data fidelity units of the unrolled network, while the other was used to define the loss.

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
N2K[86]	Adaptive Self-Supervised, Dilated convolutional J invariant network	Dec 2020	AWGN image denoising, speckle image denoising, Salt and Paper image denoising, Highly corrupted by unconventional noisy image denoising.	It addresses drawback of (predicting different brightness for salt-and-paper noise leads to appearing brightness shifting artifact always in the case of highly corrupted non-zero mean noise) of S2S and is first fully blind denoising method to clean brightness shifting got highly corrupted images without noise statistics and clean-noisy pairs. It proposes a dilated CNN using kernel-based training.

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
Nei2Nei[87]	Self-Supervised, U-Net	RRG-GAN Jan 2021	Gaussian image denoising, Poisson image denoising, Real noisy image denoising	It assumes neighboring pixels as different noisy realizations of same signal. By generating sub-sampled paired images with random neighbor sub-samplers as training image pairs and using the self-supervised training scheme with a regularization term. It is dataset-based method. Its scheme for approximating neighbor pixels leads to over smoothing, and sub sampling destroys the structural continuity.
R2R[88]	unsupervised, DnCNN	2021	Gaussian image denoising, Real noisy image denoising	It unpairs noisy images and proved that statistically training scheme has the same loss function as that of the supervised training up to a constant.

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
Noise2Score[89]	Self-Supervised, multiple trained U-Net	Jun 2021	Gaussian image denoising, Poisson image denoising, Gamma image denoising, Real noisy image denoising	It is closely related to Noisier2Noise. It requires prior knowledge of noise distribution. So when the underlying noise distribution is unknown, Noise2Score provide sub-optimal performance. This method is composed of two steps: the score function estimation by minimizing the amortized residual denoising autoencoder (AR-DAE) loss function, after which the final denoised image is obtained by applying Tweedie's formulas.
Noise2Fast[90]	Self-Supervised, U-Net	Aug 2021	Gaussian image denoising, Fluorescence microscopy image denoising	It is fast single image blind denoiser inspired by Nei2Nei, it builds four small discrete image training set using unique checkerboard downsampling for quick converge.

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
Idea-net[91]	Unsupervised, NN	Feb 2022	Gaussian image denoising, Real noisy image denoising	The key concept is to minimizing prediction variance reduction and high computation cost (especially time complexity) of existing denosier
AP-BSN[92]	Self-Supervised, U-Net like Lightweight NBSN	Mar 2022	Real noisy image denoising	It utilizes pixel-shuffle downsampling to disrupt the spatial correlation of real-world noisy images and employed asymmetric PD stride factors for training and inference. It suffers from various limitations global structure destruction, high-frequency detail loss caused by aliasing effect, and time-consuming inference process.

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
Blind2Unblind[93]	Self-Supervised, U-Net with global-aware mask mapper and blind-spot strategy	Mar 2022	Gaussian image denoising, Poisson image denoising, Real noisy image denoising	It does transition from blind to non-blind and enables blind spot schemes to use complete noisy images and overcome information loss in blindspot driven denoising. This method which is motived from classical blindspot. uses an artificial masking scheme to form blind-noisy pairs. It consists of two main components: global-aware mask mapper and re-visible loss. The mask mapper performs global denoising on blind spots and introduces regularized re-visible loss to convert the blind spots into visible ones and overcome the challenge of having less information based on blind spots. It makes blind spots visible without sub sample, noise model priors and identity mapping.

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
Poison2Sparse[94]	Self-Supervised ,NN having Encoder-decoder with sparsity and dictionary learning	Jun 2022	Poisson image denoising, Fluorescence microscopy image denoising, Monte-Carlo image denoising	It proposes a novel method for single image Poisson denoising which leverages a neural network's inductive bias. It utilizes the patch recurrence property in an image to learn a dictionary that be utilized to generate a sparse representation such that the clean image can be recovered. The dictionary elements ideally represent a set of over complete basis functions that can well represent the image. The sparse representation is then used to reconstruct the image using the learned dictionary to suppress the noise present in the input image.

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
Decay2Distill[95]	Self-Supervised, DnCNN	Aug 2022	Gaussian image denoising, Poisson image denoising.	It takes advantages of image denoising through bit compression and row shifting and allows the CNN to learn the hidden representations of an image by minimizing the loss function.
Noise2SR[96]	Self-Supervised, U-Net	Sep 2022	fluorescence microscopy image denoising, Image super-resolution.	It proposes a sub-sampler and an SR module in the denoising network, assuring that Noise2SR can be trained with constructed paired noisy images with different image dimensions and builds up a larger blind region for training the denoising network compared with other existing blind-spot networks. The weaknesses are insufficient rational behind self-supervision and subsampling technique, no ablation study, not the best result (N2N is better)

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
N2V2[97]	Self-Supervised, Customized U-Net	Nov 2022	Gaussian image denoising, Fluorescence microscopy image denoising	It solves checkerboard artifacts introduced by N2V using non-Residual U-Net using Blur Pool eliminating skip connections at the uppermost level It proposes mean and median replacement strategies as blind spot replacement excluding center pixel to remove full residual information.
DAP[98]	Unsupervised, Synergy of U-Net with Transformer concept	Feb 2023	Image Inpainting, Low vision image denoising	It incorporates self-similarity prior into the network architecture by self-attention mechanism.
AT-BSN[99]	Self-Supervised, U-Net, BSN	Mar 2023	Real noisy image denoising	It resolves the limitation of AP-BSN while minimizing the disruption of global structure and high-frequency information, thus preserving fine-grained details.

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
C-BSN[100]	Self-Supervised,U-Net	Apr 2023	Real noisy image denoising	It conditionally controls its blindness by switching the masked convolution operations. It imposes the loss on randomly downsampled sub-image on that correlation of the noise is weakened without inducing visual artifacts.
MM-BSN[101]	Self-Supervised, BSN	Apr 2023	Real noisy image denoising	It enhances BSN that can integrate the feature extracted by multi-masked convolution kernels, cannot the model size growth and preserve image details when denoising.

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
ZS-N2N[102]	Self-Supervised, Simple two-layer image-to-image CNN	Mar 2023	Gaussian image denoising, Poisson image denoising, Fluorescence microscopy image denoising	It extends N2N and Nei2Nei. It does not utilize prior information about on noise distribution or level. It assumes that nearby pixels of clean image are highly correlated, while noise independent and unstructured. It convolves the noisy test image with two fixed filters to yields two downsampled images (approximation of two noisy observations of the same clean image) And then train a lightweight network with regularization to map one downsampled image to the other.

Table 2.3: The denoising models trained on single/multiple image pairs (without clean)

Model Name	**Training model with Network**	**First Proposed on**	**Area of Applications**	**Remarks**
SwinIA[103]	Self-Supervised, Swin Transformer	May 2023	Gaussian image denoising, Image super-resolution, JPEG compression artifact reduction	It is Swin Transformer-based Image Autoencoder (SwinIA), the first convolution-free architecture for self-supervised denoising. It is trained as a plain autoencoder by minimizing the mean squared error (MSE) computed over the full image. It does not use clean images, pre-training, knowledge distillation, input masking, auxiliary regularization losses, multiple forward passed.
PI-DDPM[104]	Self-supervised, Denoising Diffusion Model Based	Jun 2023	Gaussian image denoising, Fluorescence microscopy image denoising	Its reconstructions show more consistency with the confocal image than the conventional DDPM or U-Net irrespective of whether the input image comes from widefield or confocal microscopy.

CHAPTER 3

Methods and Materials

In this chapter, the thesis discusses the methods and material used in denoising of high-resolution confocal microscopy images of dendritic spines using Noise2Void algorithm.

3.1 Model Architecture for Noise2Void

NOISE2VOID (N2V) [1], a denoising self-supervised learning method that is motivated by Noise2Noise (N2N) [67] model using independent pairs of noisy images, that gets trained directly on the body of data to be denoised. It does not require noisy image pairs, nor clean target images. This model is built on two assumptions:

- Actual part of the Image has structure but noise does not structure.
- Noise can not be predicted using neighbourhood pixels, whereas the image structure can be predicted by looking into neighbourhood pixels.

The conventional network (CNN) degenerates because of simply learning the identity when the training is conducted using the same noisy image as the input and target. N2V model relies on blind spot network Figure 3.1 which is key feature. The working principle of Noise2void model is as follows : In N2V, patches that segmented images are inputted to the CNN. During training, for a image patch, the central pixel is included as part of target and prediction in the Conventional network's receptive field, whereas central pixel is excluded in Noise2Void's receptive field. Instead of dropping of central pixel of image patch by making as 0, central pixel is replaced by a randomly selected any other pixel from the same patch and then the modified image patch becomes as input to CNN during training. The original input patch with unmodified pixel values becomes target image

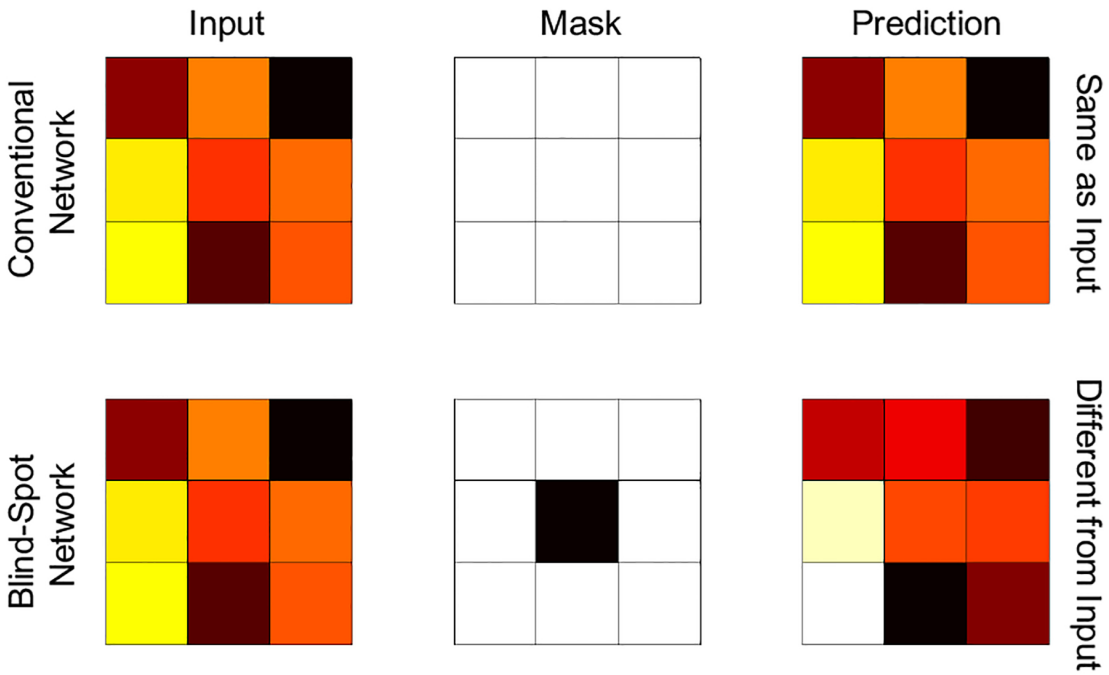


Figure 3.1: The blind spot concept : Unlike a conventional network, a blind-spot network has a masked receptive field that excludes the central pixel and can learn to suppress noise by focusing on the neighboring pixels. Thus, it can generate a prediction distinct from the input even when the input and target images are identical and noisy.

Figure Courtesy : [105]

patch. The training loss is minimized between modified image patch and original image patch as part of learning process.

N2V allows network training without requiring a large number of data because N2V is a self-supervised learning method, one of the N2V merits. Although enormous training data provides a robust DnCNN model, in cases of a small number of training data, the deterioration of the predicted image is caused by the influence of training images with various structures and textures. Therefore, the images in this study augmented the application of noise reduction, and the network training was conducted with the augmented data.

Advantages of Noise2Void denoising algorithm :

- Other classic deep-learning models for denoising, such as noise-to-low noise and noise-to-noise, may not be usable if the required outputs cannot be acquired. For N2V model, each pixel is replaced with another pixel in the same image.
- Training N2V models usually does not require many epochs.
- N2V models are very good for reducing salt and pepper noise.

Training Inputs/Parameters :

1. **Input images :**
    - Single or list of input images based on which model will be trained.
    - Supported file formats: OME-TIFF, CZI and TIF. Internally CZI and TIF images are converted to OME-TIFF.
    - The module can work with time-series, Z stack and multichannel images also. Internally, each 2D plane of an image file is processed separately.
    - OME-TIFF images must be 5 dimensional.
    - Images with multiple scenes are also supported.
    - The bit depth across input images should be unified. It is not recommended to have multiple bit depths present in the input images.
2. **Training Sources Folder :**
    - These is the path to folders containing the noisy images to be used for training
3. **Model Name :**
    - This is to provide model name eg. Use only my_model -style., not my-model (Use "_" not "-"). The space(s) are not allowed in the name. If the name of an existing model (saved in the same folder) is given, then it will be overwritten.
4. **Model Path :**
    - The path where created N2V model will be saved once trained.
5. **Number of Epochs :**
    - Maximum number of epoch used for model training. Early stopping might shorten the training time.
    - Default value: 100
    - Range: 0 - 1000
6. **Patch size (pixels) :**
    - Noise2Void divides the image into patches for training. Input the size of the patches (length of a side). The value should be between 64 and the dimensions of the image and divisible by 8.

- Default value: 64

7. **Neighborhood Radius :**
   - Radius around target pixel from which the model chooses random one for replacement.
   - Default value: 5
   - Range: 1 - 10
8. **Training Loss Method :**
   - Options are mae or mse
   - Default value: mae
9. **Batch Size :**
   - This parameter defines the number of patches seen in each training step. Noise2Void requires a large batch size for stable training. Reduce this parameter if your GPU runs out of memory.
   - Default value: 128
10. **Number of Steps :**
    - It defines the number of training steps by epoch. By default this parameter is calculated so that each image / patch is seen at least once per epoch.
    - Number of patch / batch_size
11. **Percentage validation :**
    - The percentage of training dataset to be usedto validate the network during the training.
    - Default value: 10
12. **Initial Learning Rate :**
    - The initial value to be used as learning rate
    - Default value: 0.0004

Output :

1. **Trained model weights :**
   - Model's best weights.
   - Format: h5
   - Contains model's structure and weights in HDF5 format.
2. **Trained model config :**
   - Model's configuration file.
   - Format: json
   - Contains various model hyperparameters.
3. **Training history :**
   - Graphical depiction of model's training and validation losses over epochs..
   - Format: png

3.2 Loss Functions and optimizer

A loss function is a function that measures how well a machine learning model's predicted outputs match the true output labels. The loss function is used to optimize the model by minimizing the loss, which means that the model makes fewer mistakes on the training data. The loss function is an important part of the machine learning process because it provides a way to evaluate the model's performance and guide the optimization process by indicating how the model's predictions differ from the true output labels. Some common loss functions in neural networks used for regression tasks include mean absolute error (MAE) loss, mean squared error (MSE) loss, mean squared logarithmic error (MSLE) loss, and Root Mean Squared Error (RMSE). RMSE and MSLE are left out of considerations for evaluation in this thesis.

3.2.1 Mean Squared Error (MSE)

The Mean Squared Error (MSE) calculates the squared error or in other words, the squared difference between the actual output and the predicted output for each sample. Sum them up and take their average. Some people call the MSE by the name of L2 Loss.

$$\text{MSE}(y, \hat{y}) = \frac{\sum_{i=0}^{n-1}(y_i - \hat{y}_i)^2}{n}$$

- $\hat{y}_i$ - Predicted Output
- y_i - Actual Output
- n - Training samples in each mini batch (if not using mini batch training, then n = Training sample). If we have 1000 training samples and we are using a batch of 100, it means we need to iterate 10 times so in each iteration there are 100 training samples so n = 100.
- No matter if it is either ($\hat{y}_i$ - y_i) or (y_i - $\hat{y}_i$), it provides the same result as it does square the distance.

The MSE loss is differentiable, meaning we can use it with gradient-based optimization algorithms. It is also a convex function with a single global minimum that optimization algorithms can find. One of the advantages of the MSE loss is that it is well-behaved and smooth, making it easier to optimize than other loss functions in neural networks. However, it is also sensitive to outliers in the data, making it less suitable for some tasks.

3.2.2 Mean Absolute Error (MAE)

Mean absolute error (MAE) loss calculates the mean of the absolute differences between predicted and true output. The MAE loss is calculated by taking the average absolute difference between the predicted and true values for all the samples in the dataset.

The following equation gives the MAE loss:

$$\text{MAE}(y, \hat{y}) = \frac{\sum_{i=0}^{n-1} |y_i - \hat{y}_i|}{n}$$

- $\hat{y}_i$ - Predicted Output
- y_i - Actual Output
- n - Training samples in each mini batch (if not using mini batch training, then n = Training sample). If we have 1000 training samples and we are using a batch of 100, it means we need to iterate 10 times so in each iteration there are 100 training samples so n = 100.
- No matter if you do ($\hat{y}_i$ - y_i) or (y_i — $\hat{y}_i$), you will get the same result because, in the end, you take the absolute distance.

3.3 Network Training, Fine-Tuning, and Validation

The network architecture of N2V is as a U-net[] with the depth of 3, kernel size of 3, batch normalization, and a linear activation function in the last layer. The network has 96 feature maps on the initial level, which get doubled while the network gets deeper. The network model was trained for 50 epochs using the loss of function *mae* as well as *mse* and a learning rate of 0.0004. Other parameters were as follows: Batch size as 128, patch size as (64x64).

The training was run on Keras (version 2.3.1) with TensorFlow (version 2.15.0) using Tesla T4 GPU.

3.4 Data Set

We used in-vitro and in-vivo image of dendritic spines for training. 2D images are obtained from 3D confocal images using MIP technique.

3.5 Metrics for quantitative and qualitative assessments of N2V denoising

In this study, the following quantitative and qualitative metrics were evaluated for the assessment of denoising performance of confocal microscopy images of dendritic spines using Noise2Void :

3.5.1 Histogram

In an image processing context, the histogram of an image normally refers to a histogram of the pixel intensity values. This histogram is a graph showing the number of pixels in an image at each different intensity value found in that image. For an 8-bit grayscale image there are 256 different possible intensities, and so the histogram will graphically display 256 numbers showing the distribution of pixels amongst those grayscale values.

An image histogram is a graph of pixel intensity (on the x-axis) versus number of pixels (on the y-axis). The x-axis has all available gray levels, and the y-axis indicates the number of pixels that have a particular gray-level value.

3.5.2 Plot Profile using Imagej Tool

It displays a two dimensional graph of intensities of pixels along a line within the image. The x axis represents distance along the line in terms of pixel and the y axis represents the pixel intensity.

3.5.3 SSIM

SSIM is an index measuring the structural similarity between two images. It is valued between -1 and 1. When two images are nearly identical, their SSIM is close to 1. Structural information is the idea that the pixels have strong inter-dependencies especially when they are spatially close. These dependencies carry important information about the structure of the objects in the visual scene. SSIM is a perception-based model that considers image degradation as perceived change in structural information, while also incorporating important perceptual phenomena, including both luminance masking and contrast masking terms. SSIM satisfies the identity of indiscernible, and symmetry properties, but not the triangle inequality or non-negativity, and thus is not a distance function.

The SSIM index is calculated on various windows of an image. The measure between two windows x and y of common size N X N is:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1) + (2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{3.1}$$

with :

- μ_x the pixel sample mean of x;
- μ_y the pixel sample mean of y;
- σ_x^2 the variance of x;
- σ_y^2 the variance of y;
- σ_{xy} the covariance of x and y;
- $c_1 = (k_1 L)^2, c_2 = (k_2 L)^2$ two variables to stabilize the division with weak denominator;
- L the dynamic range of the pixel-values (typically this is $2^{\#bits\ per\ pixel} - 1$);
- $k_1 = 0.01$ and $k_2 = 0.03$ by default.

The SSIM formula is based on three comparison measurements between the samples of x and y : luminance (l), contrast (c) and structure (s). The individual comparison functions are

$$l(x,y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \tag{3.2}$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \tag{3.3}$$

$$s(x,y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3} \tag{3.4}$$

with, in addition to above definitions:

- $c_3 = c_2/2$

SSIM is then a weighted combination of those comparative measures:

$$SSIM(x,y) = l(x,y)^\alpha \cdot c(x,y)^\beta \cdot s(x,y)^\gamma \tag{3.5}$$

Setting the weights α, β, γ to 1, the formula can be reduced to the form shown above.

3.5.4 NR-IQA

This is a challenging task to predict denoising quality metrics like Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) when no ground-truth image is available. No-reference image quality metrics can be applied to image denoising quality assessment. In the literature, the NR-IQA model is designed based on two perceptions, such as natural scene statistics (NSS)[106] and the human visual system (HVS). In the NSS approach, the features are assumed to be statistically regular. It is considered to vary with the levels of distortions. On the contrary, HVS-based models use perceptual features to mimic the visual content based on human vision physiology . The natural scene statistics include BRISQUE[107] and NIQE[108] both of which are no-reference, since they do not require any reference picture to measure quality against. BRISQUE is one of the most commonly used methods for NR-IQA. It uses only spatial features (without any frequency domain features). This method looks at image quality in a more holistic way, meaning it provides a single score for the entire image's quality, and not sub-scores for each image quality category. BRISQUE is used for NR-IQA assessment in this thesis.

3.6 Datasets

2D images of dendritic spines for invitro and invivo which were used as datasets, were obtained by converting 3D invitro and invivo confocal microscopy images

to 2D via maximum intensity projection technique using ImageJ tools. In this experiment, two type of images such as invitro and invivo images of dendritic spines were used. Invivo images are confocal microscopy images of dendritic spines of real-time living mouse, whereas invitro images are not of real-time time, rather is of dead mouse.

- **Training Datasets**

  The training datasets are divided into two categories: invitro-noisy and invivo-noisy images OF 2D. 32 invitro images having dimension of 512 X 512 Pixels and 20 invivo images having dimension of 1024 X 1024 pixels were used for training.

- **Validation Datasets**

  For validation or evaluation, 100 images of invitro-noisy, 5 images of invivo-noisy and 10 images of invitro having manually added noise via ImageJ. As no ground truth or clean image was available, so denoising performance of Noise2Void for confocal microscopy image of dendritic spines were conducted using noisy images only.

CHAPTER 4

Results and Discussion

In this chapter, we evaluate the denoising performance of Noise2Void model for confocal microscopy image of dendritic spines.

4.1 Noise2Void 2D Trained Models

To evaluate the performance, total four trained models using Noise2Void were created using combination of image type and loss functions. The details of trained models are given in Table 4.1. Both invivo and invitro images together were not used for creating trained model as both available images are having different resolutions.

Table 4.1: Details of Trained Models

Training Parameters	**Trained Model1**	**Trained Model2**	**Trained Model3**	**Trained Model4**
Loss Function	mae	mse	mae	mse
Image Type	invitro	invitro	invivo	invivo
Image Dimensions (Pixels)	512 X 512	512 X 512	1024 X 1024	1024 X 1024
Patch Size	64 X 64	64 X 64	64 X 64	64 X 64
Batch Size	128	128	128	128
Number of Steps	116	116	289	289
Number of patches	16384	16384	40960	40960
Number of epochs	50	50	50	50
% of validation	10	10	10	10
initial learning rate	0.0004	0.0004	0.0004	0.0004

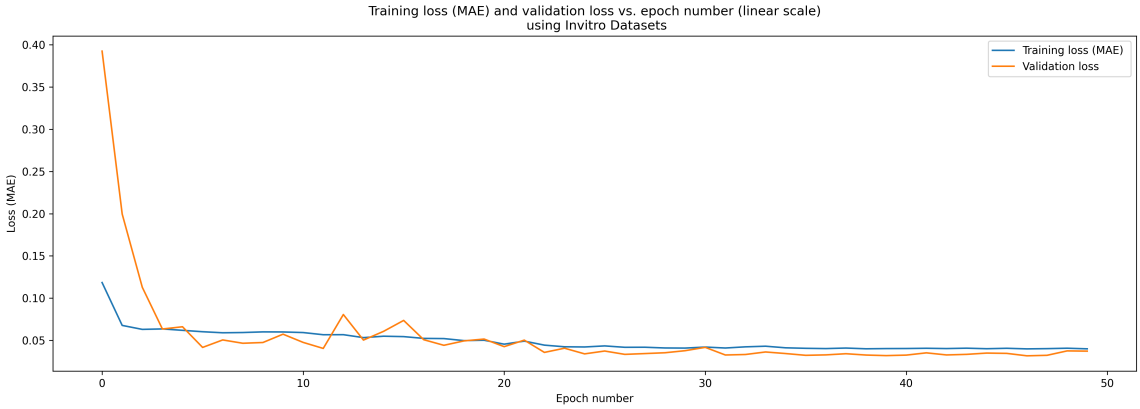


Figure 4.1: Training loss (MAE) and validation loss vs epoch number (linear scale) using Invitro Datasets

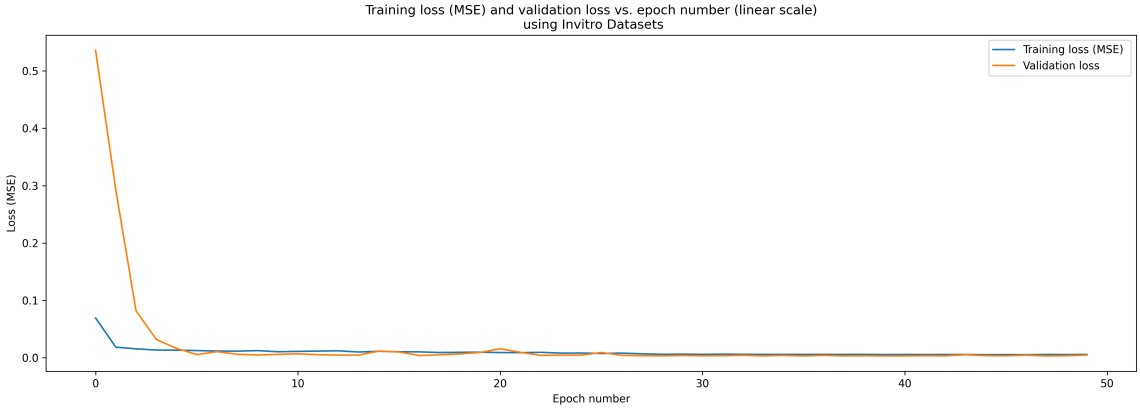


Figure 4.2: Training loss (MSE) and validation loss vs epoch number (linear scale) using Invitro Datasets

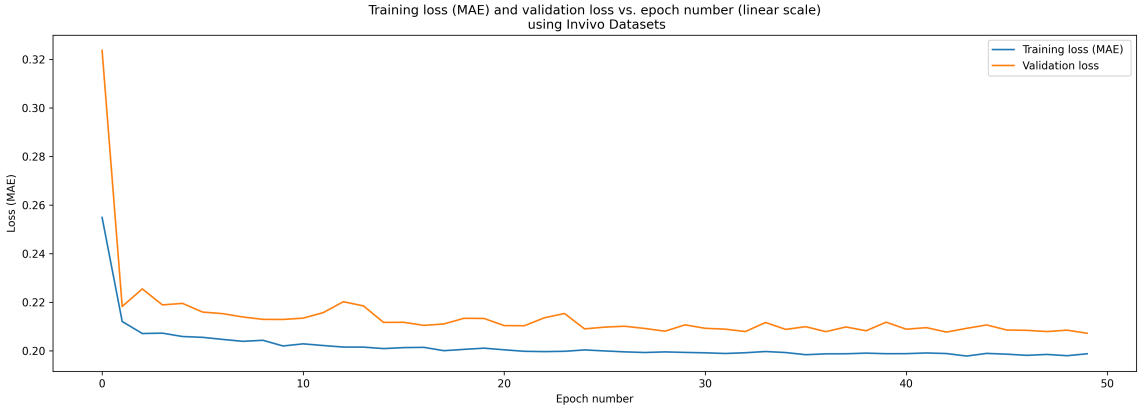


Figure 4.3: Training loss (MAE) and validation loss vs epoch number (linear scale) using Invivo Datasets

The loss and validation function vs epoch numbers for Trained model 1,2,3 and 4 are shown in Figure 4.1, Figure 4.2, Figure 4.3 and Figure 4.4 respectively.

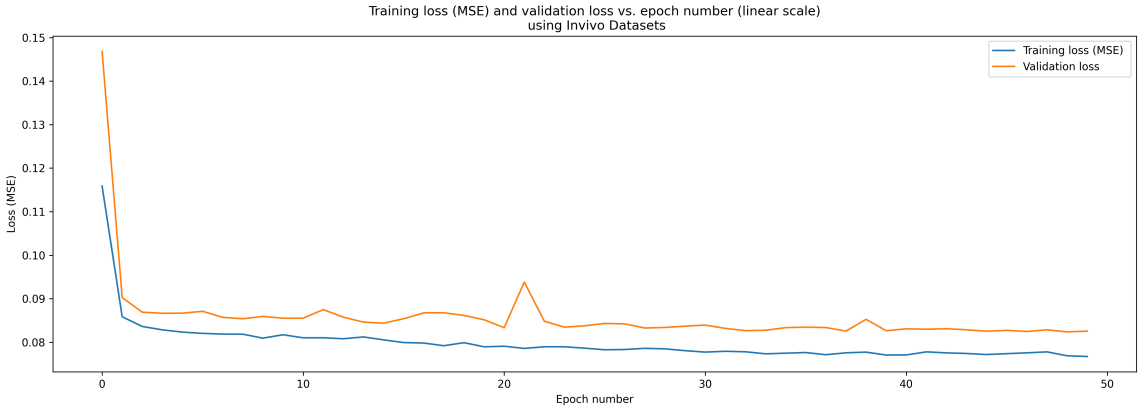


Figure 4.4: Training loss (MSE) and validation loss vs epoch number (linear scale) using Invivo Datasets

4.2 Experimental Results

All four models created using Noise2Void 2D model, were evaluated for both unseen noisy datasets (invivo images, invitro images and invitro images having noise added manually which were not used for training) as well as seen noisy datasets(invivo, invitro images which were used during training).

The following quantitative and qualitative performance metrics were used for evaluation of denoising performance :

- No reference image quality analysis (NR-IQA) score for each image is shown. The lower value shows good quality.
- SSIM score of pre-denoised image vs post-denoised image is shown. Closure to 1 means the lower structural change during denoising.
- SSIM score of pre-denoised image vs post-denoised image is shown to depict the structural change visually.
- The histogram of pre-denoising and post-denoising is plotted to show the change of intensity distribution.
- The line plot profile for yellow line drawn on each image is shown to illustrate intensity changes along the line.
- Sub-image marked using red rectangular box of each pre-denoising and post-denoising is zoomed at below to visualise the effect denoising across various models used.

In this experimental results, no ground truth was used to validate the performance of Noise2Void due to non-availability of ground truth.

The visual comparison of denoising performance of unseen invitro image of dendritic spines are shown by zooming of sub-image [(340, 315) to (380,335)] is shown in FigureFigure 4.5 and Figure 4.6.

The visual comparison of denoising performance of unseen invivo image of dendritic spines are shown by zooming of sub-image [(50, 250) to (90,270)] is shown in Figure 4.7 and Figure 4.8.

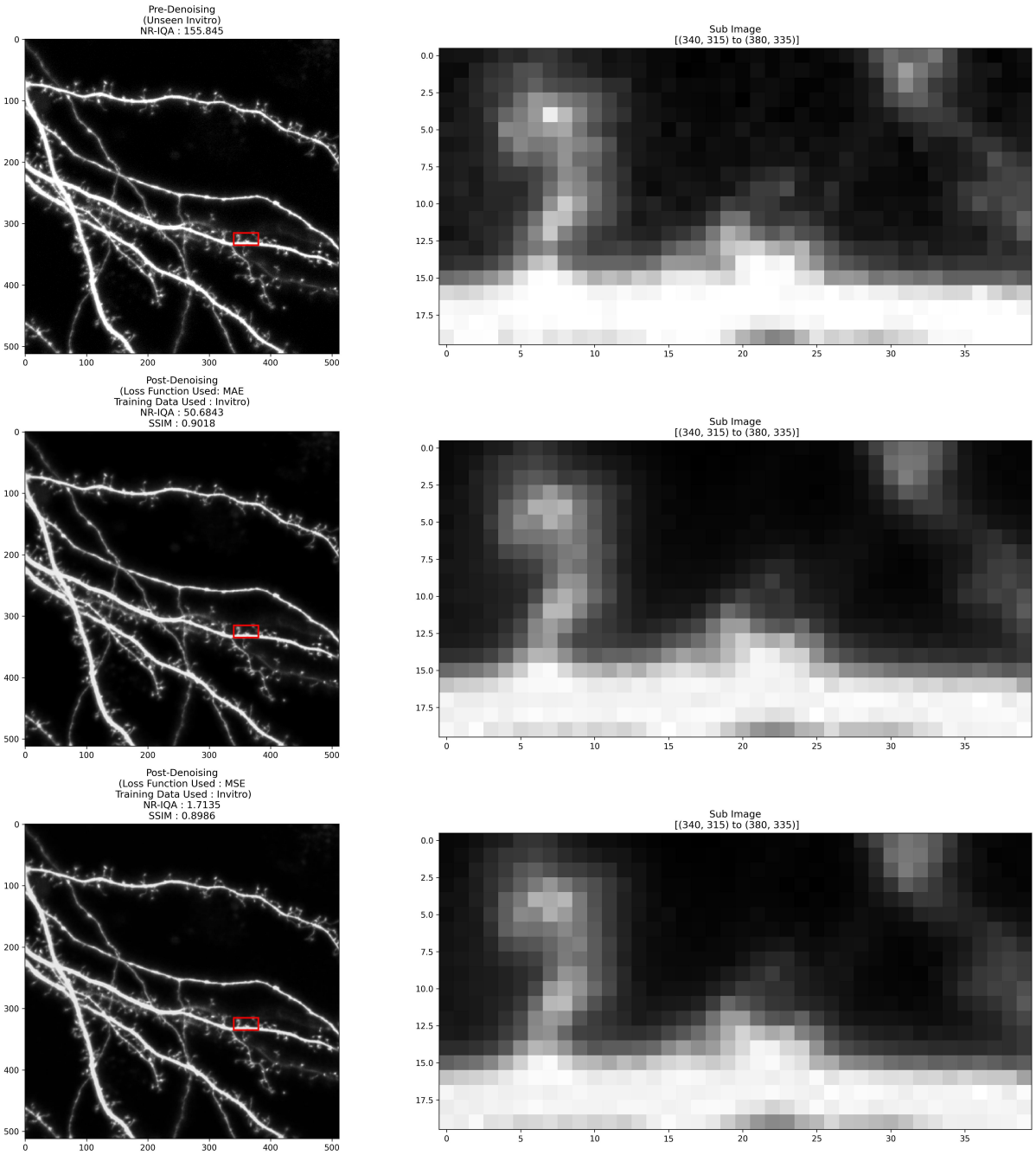


Figure 4.5: Visual denoising evaluation of unseen invitro image of dendritic spines using model trained with invitro images for both mae and mse loss functions

The comparison of performance of Noise2Void 2D Training Models for **unseen**

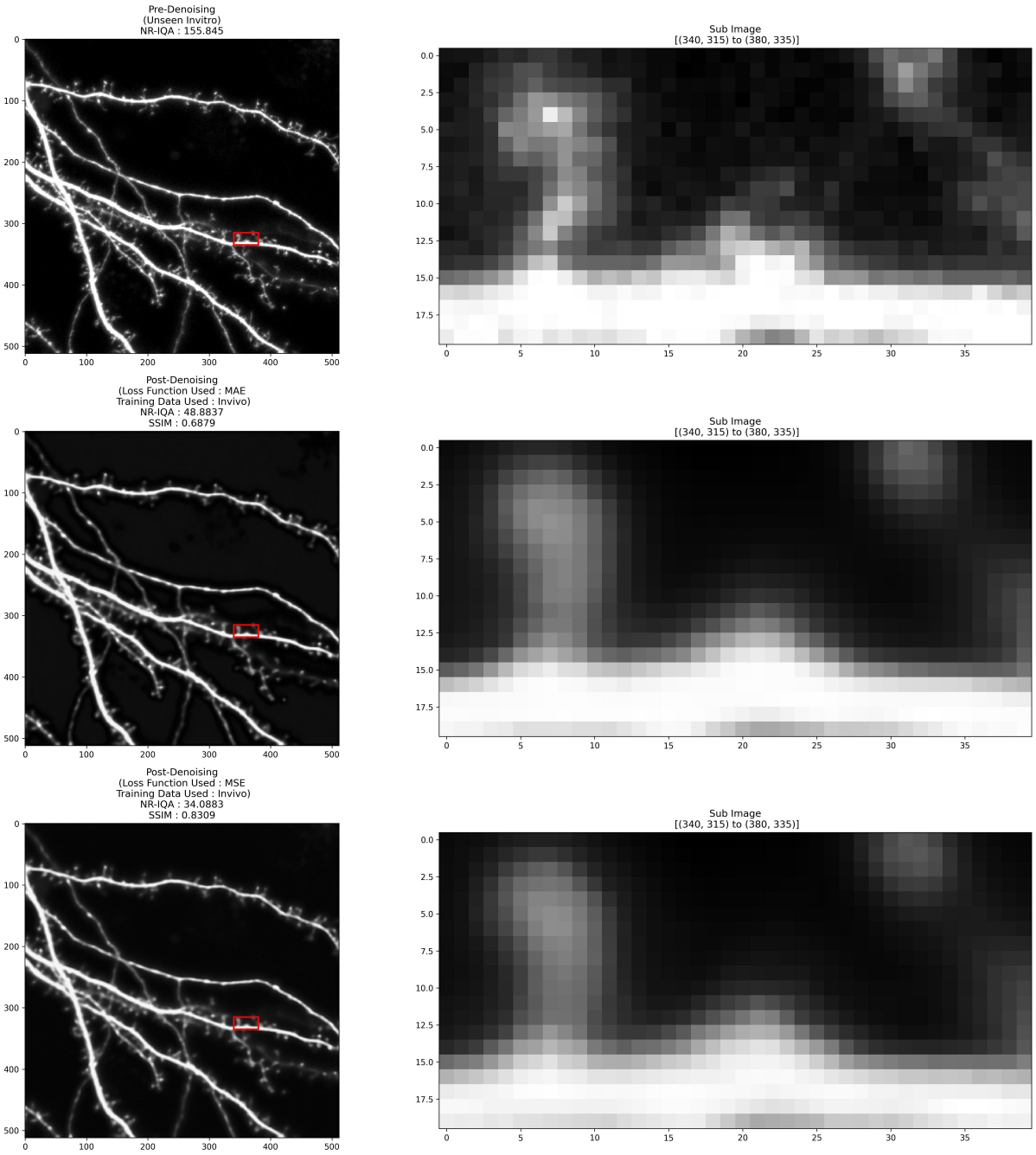


Figure 4.6: Visual denoising evaluation of unseen invitro image of dendritic spines using model trained with invivo images for both mae and mse loss functions

datasets for invitro, invivo and invitro with added noise were denoised through all fours trained models are shown in Figure 4.9, Figure 4.10, Figure 4.11 respectively.

Similarly, the comparison of performance of Noise2Void 2D Training Models for **seen datasets** for invitro and invivo were denoised through all fours trained models are shown in Figure 4.12, Figure 4.13 respectively.

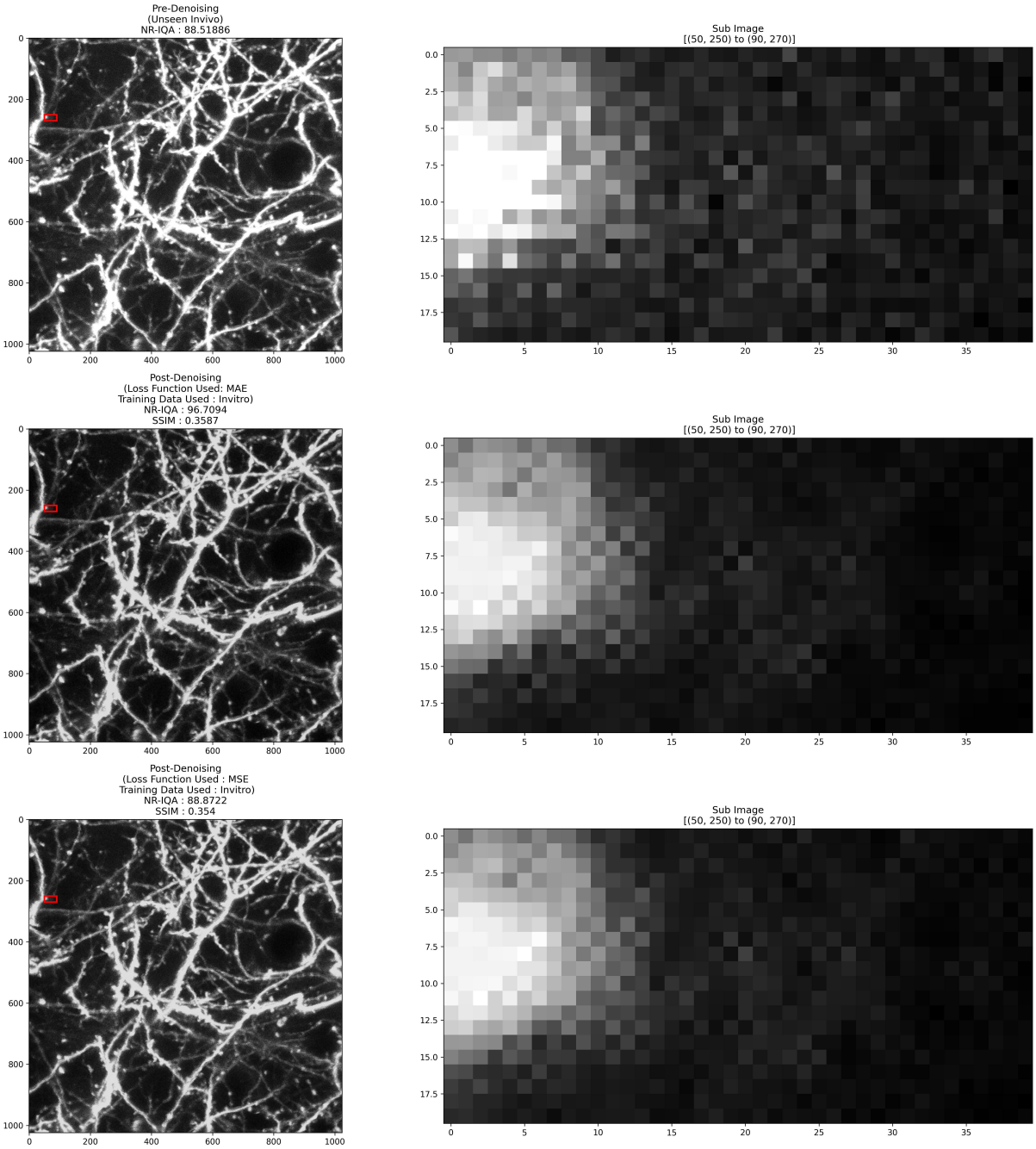


Figure 4.7: Visual denoising evaluation of unseen invivo image of dendritic spines using model trained with invitro images for both mae and mse loss functions

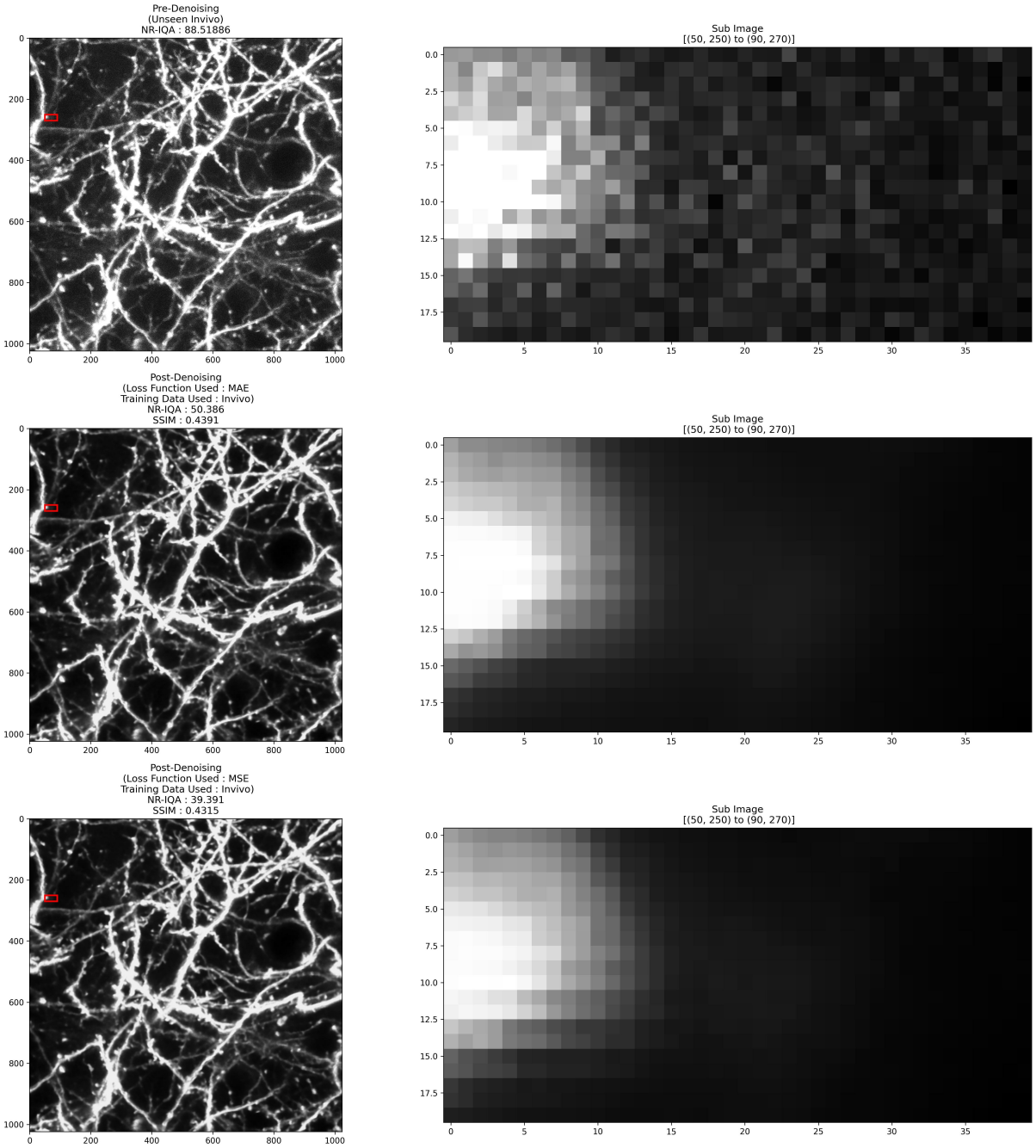


Figure 4.8: Visual denoising evaluation of unseen invivo image of dendritic spines using model trained with invivo images for both mae and mse loss functions

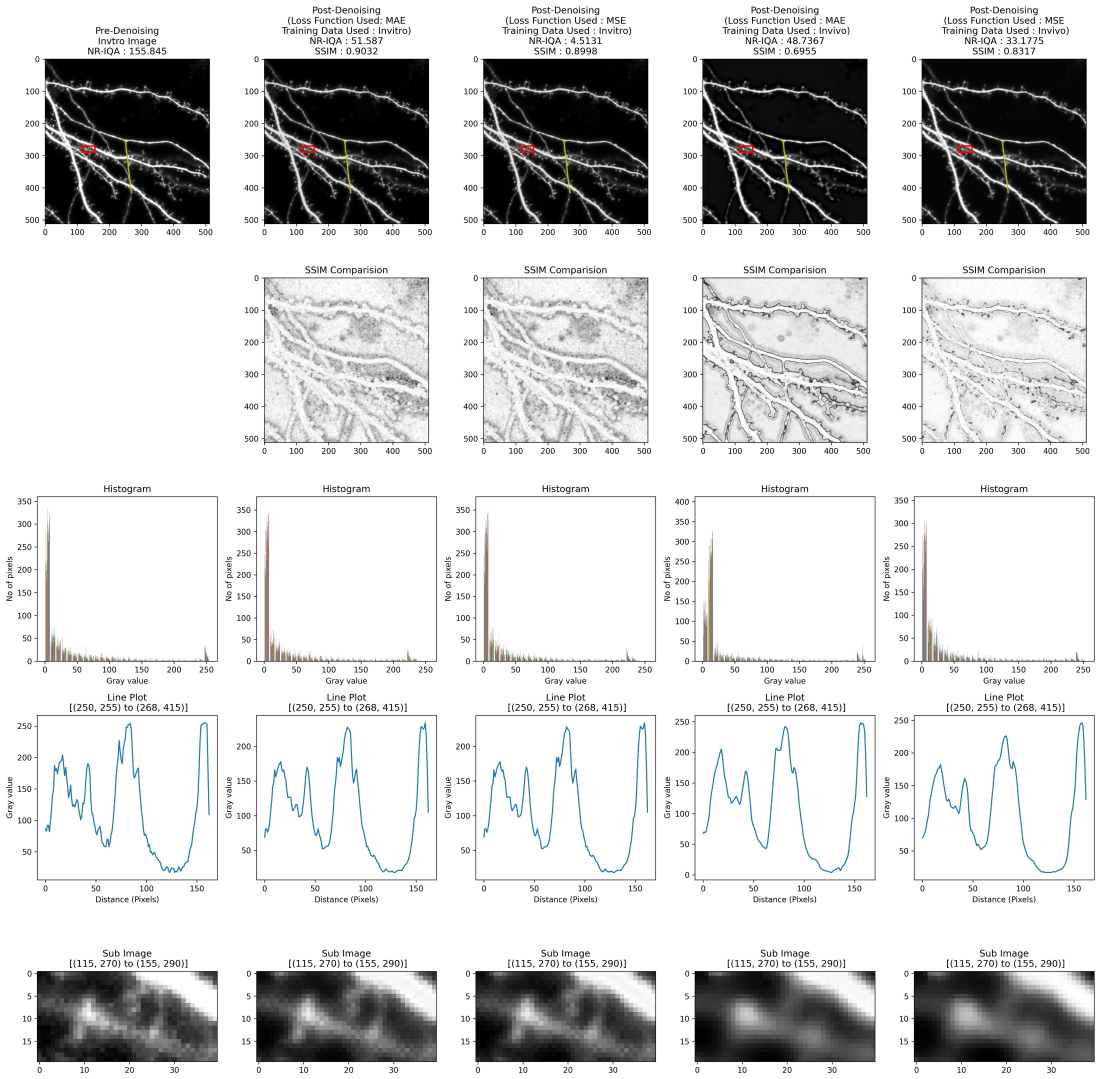


Figure 4.9: Denoising comparison Of Invitro Image of Dendritic Spines Not Used In Model Training

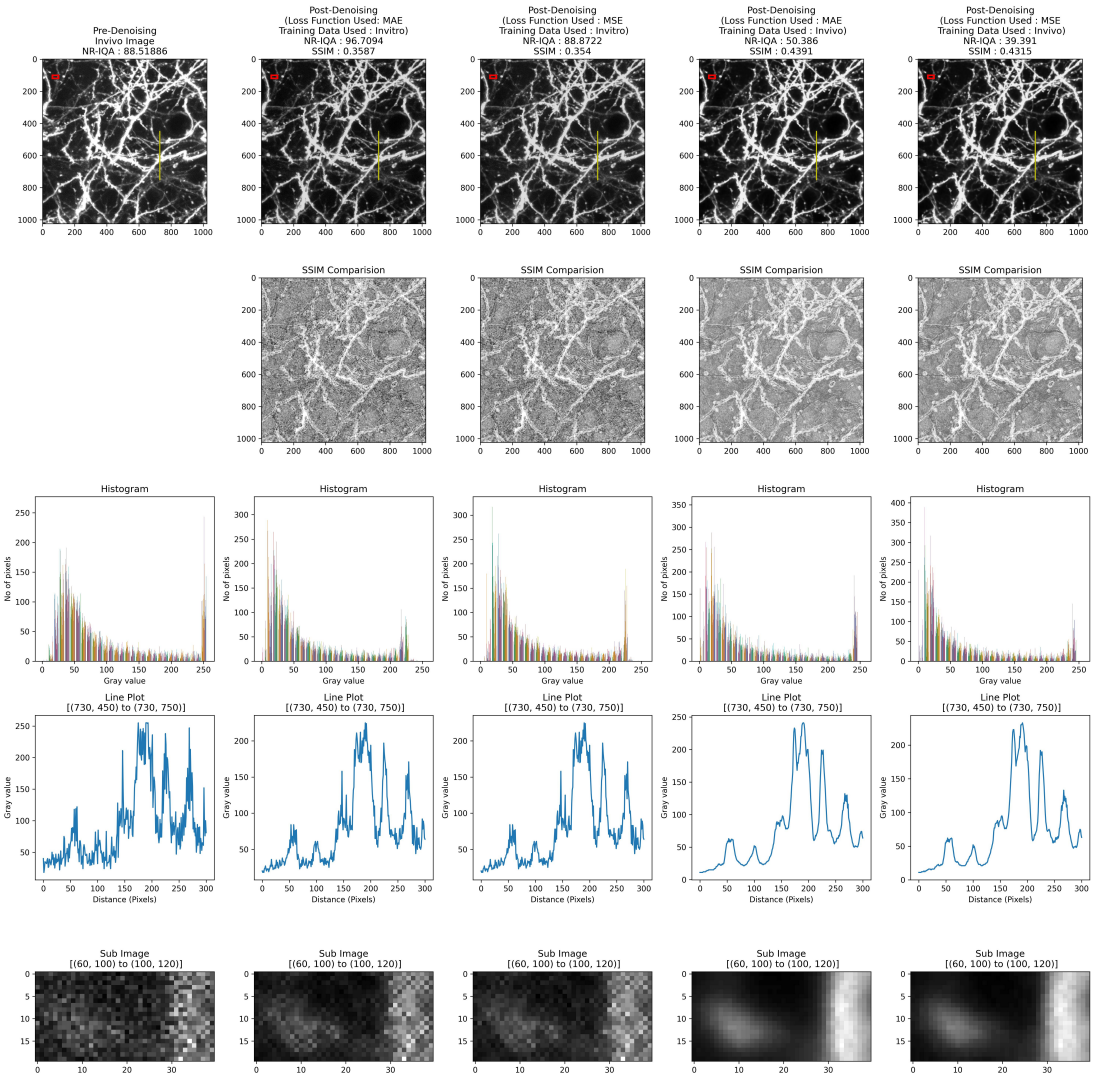


Figure 4.10: Denoising comparison of Invivo Image of Dendritic Spines Not Used In Model Training

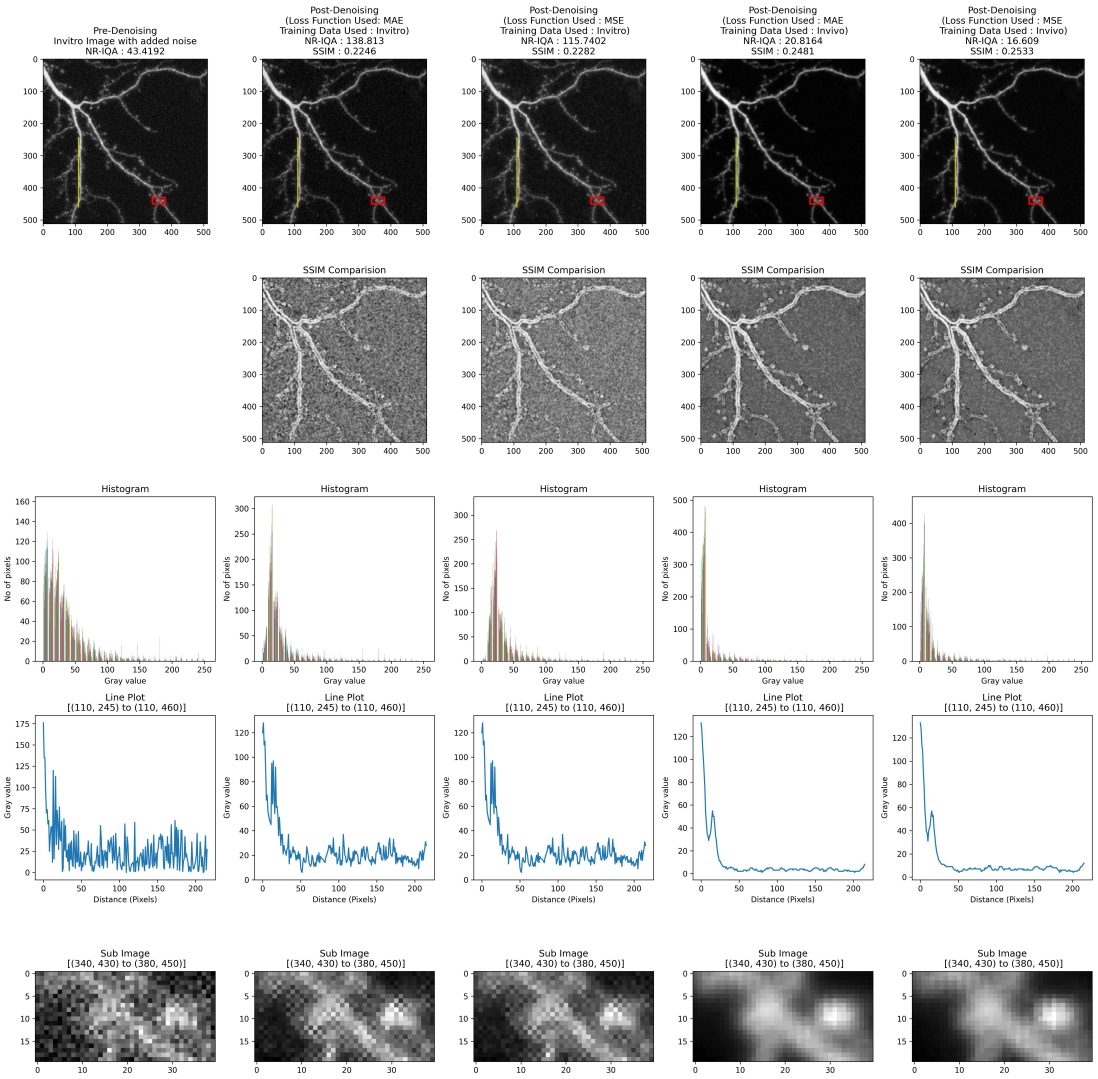


Figure 4.11: Denoising comparison of Invitro Image of Dendritic Spines Having Added Noise Not Used In Model Training

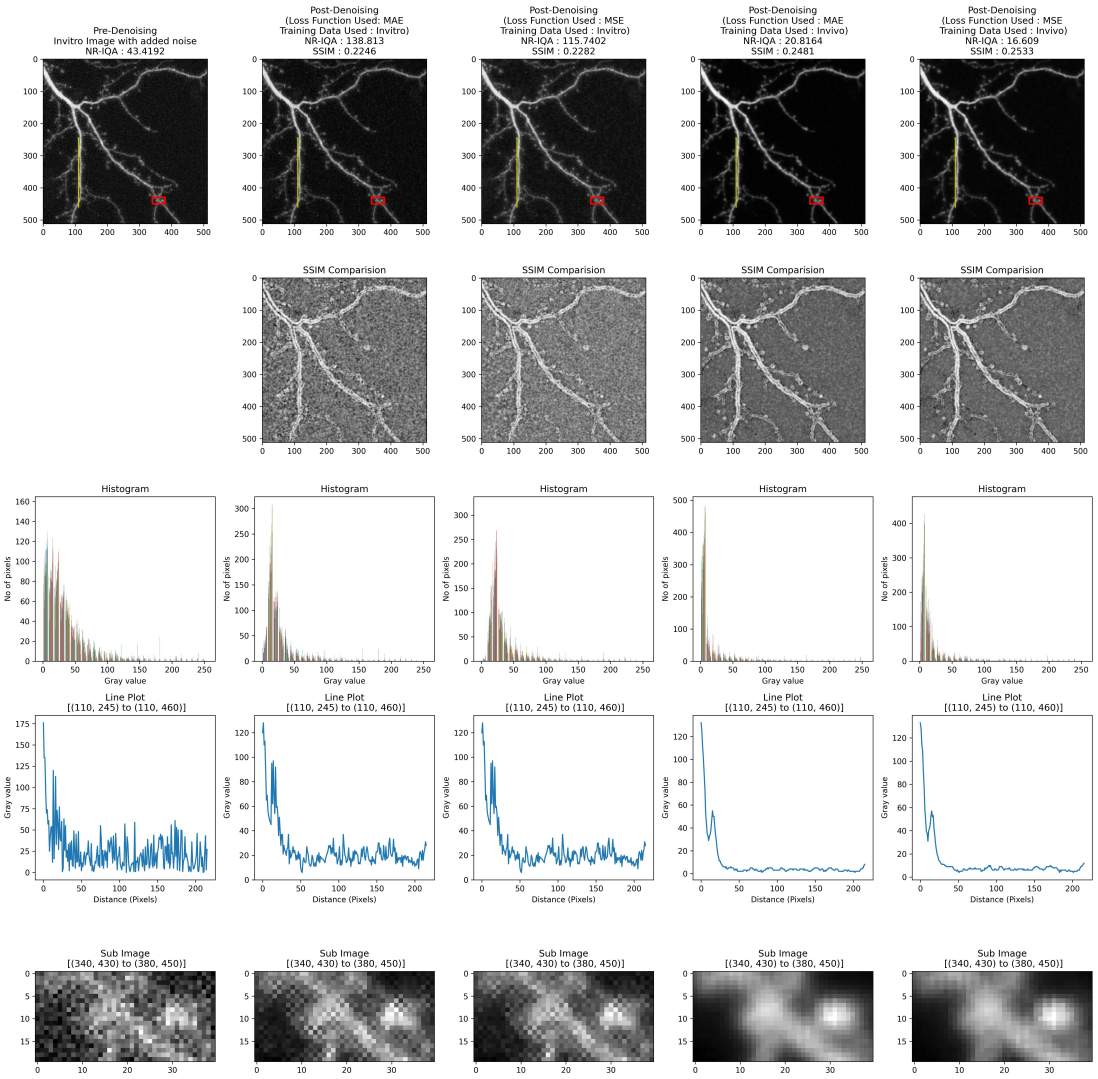


Figure 4.12: Denoising comparison of Invitro Image of Dendritic Spines Used In Model Training

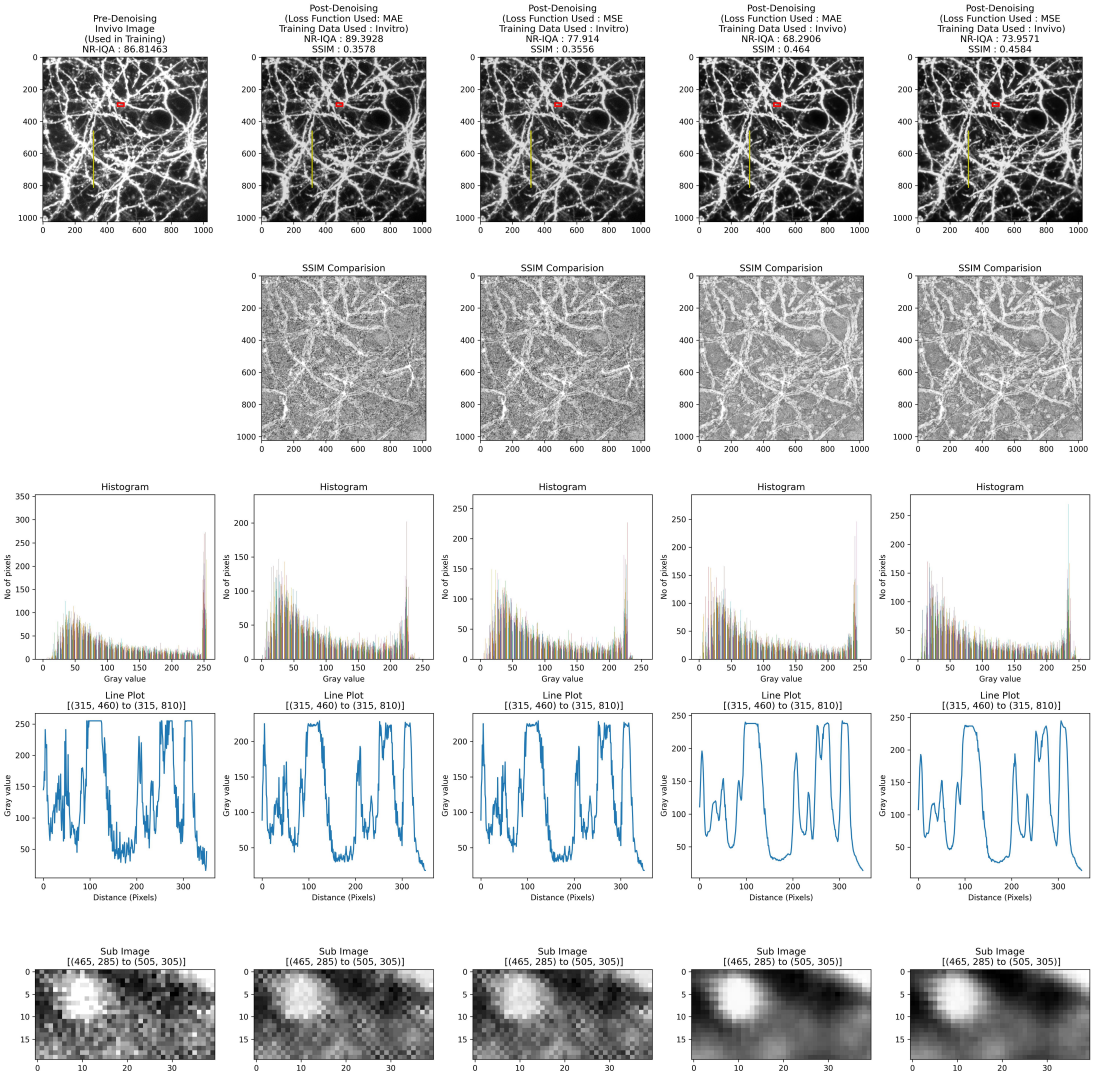


Figure 4.13: Denoising comparison of Invivo Image of Dendritic Spines Used In Model Training

CHAPTER 5

Conclusions

In this thesis, this experimentation of using Noise2Void algorithm for confocal microscopy images of dendritic spines was conducted on two distinct images : ivitro and invivo images. It is observed that the invivo image of dendritic spine are more noisier than invitro images, as invivo images are taken real-time living mouse. Various qualitative and quantitative metrics are used for evaluation. With respect to loss function, we observe that NR-IQA metric obtained using trained model 2, 4 (MSE loss function used) shows much better in comparison with trained model 1, 3 (MAE loss function used). With respect to training data used, it is clearly visible that the trained model 3, 4 which are trained using invivo image effective and efficient denoising performance for invivo as well as invitro, whereas trained model 1,2 are not showing effective and efficient denoising performance. The SSIM metrics of noised vs denoised image also show that Noise2Void model is able to remove noises in confocal microscopy of dendritic spines. The line plot profile and sub image zooming shows visually that Noise2Void has reduced the noise and introduce smoothness as well.

As Noise2Void is self-supervised, ground truth/clean images are not needed.

In this experiment, the performance of Noise2Void model using noisy images only is validated and it is clearly seen that Noise2Void can reduce the noises. It can be concluded that Noise2Void algorithm can be applied for confocal microscopy images of dendritic spines for denoising. However, further experiment, studies and enhancement are required for quality improvement of denoising of confocal microscopy images of dendritic spines.

5.1 Scope for Future Works

It is observed that Noise2Void algorithm is able to denoise confocal microscopy images of dendritic spines, so there is scope of improvement for Noise2Void in

future for efficient image denoising of confocal microscopy images of dendritic spines. By default, Noise2void uses either MAE or MSE as loss function, there is scope to explore this model using other loss function by implementing such as Root Mean Squared Error (RMSE), Mean Squared Logarithmic Error Loss (MSLE), or any suitable loss functions.

Noise2Void model was trained using either invitro or invivo images, but not both images together. Since, invivo and invitro images with same dimensions were not available, much exploration of model training using both image types could not be performed. However, in this experiment, it was tried to train the model using both images together having different dimensions, but training graph of the training loss and validation loss vs epoch count did not show promising. So it is needed to train Noise2Void model using both invitro and invivo images together for further analysis.

Unfortunately, the blind-spot training underlying Noise2Void can lead to rather visible checkerboard artifacts, thereby reducing the quality of final predictions considerably. It is observed few unwanted patches being developed while denoising of dendritic spines, also sharpness of the image is reduced which is observed specially in trained model4 (invivo as training data and mse as loss function). Already, further enhancements were done to fix few issues of Noise2Void in N2V2[97] and Probabilistic Noise2Void[77]. So, in future, more further studies, experiments and enhancement on Noise2Void model are needed to improve denoising performance of confocal microscopy images of dendritic spines.

Bibliography

[1] A. Krull, T.-O. Buchholz, and F. Jug, "Noise2void - learning denoising from single noisy images," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2124–2132, 2018.

[2] *Confocal microscopy*, `https://www.edmundoptics.com/knowledge-center/application-notes/microscopy/confocal-microscopy//`.

[3] S. Ramón and S. Cajal, "Textura del sistema nervioso del hombre y de los vertebrados," *Imprenta y Librería de Nicolás Moya, Madrid (1899–1904)*, 1904.

[4] W. Wu, M. Chen, Y. Xiang, Y. Zhang, and Y. Yang, "Recent progress in image denoising: A training strategy perspective," *IET Image Processing*, vol. 17, no. 6, pp. 1627–1657, 2023.

[5] 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)*, Ieee, vol. 2, 2005, pp. 60–65.

[6] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *IEEE Transactions on image processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[7] F. MARTIND *et al.*, "Adatabaseof humansegmentednaturalimagesanditsapplicationto evaluating segmentation algorithms and measuring ecologicalstatistics," *IEEEInternationalconference oncomputerVision. Vancouver*, vol. 416, p. 423, 2001.

[8] S. Roth and M. J. Black, "Fields of experts," *International Journal of Computer Vision*, vol. 82, pp. 205–229, 2009.

[9] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *2009 IEEE 12th international conference on computer vision*, IEEE, 2009, pp. 2272–2279.

[10] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," in *2011 international conference on computer vision*, IEEE, 2011, pp. 479–486.

[11] H. C. Burger, C. J. Schuler, and S. Harmeling, "Image denoising: Can plain neural networks compete with bm3d?" In *2012 IEEE conference on computer vision and pattern recognition*, IEEE, 2012, pp. 2392–2399.

[12] U. Schmidt and S. Roth, "Shrinkage fields for effective image restoration," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 2774–2781.

[13] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 2862–2869.

[14] Y. Chen and T. Pock, "Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 6, pp. 1256–1272, 2016.

[15] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," *IEEE transactions on image processing*, vol. 26, no. 7, pp. 3142–3155, 2017.

[16] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep cnn denoiser prior for image restoration," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 3929–3938.

[17] S. Lefkimmiatis, "Non-local color image denoising with convolutional neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 3587–3596.

[18] P. Liu, H. Zhang, K. Zhang, L. Lin, and W. Zuo, "Multi-level wavelet-cnn for image restoration," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2018, pp. 773–782.

[19] T. Plötz and S. Roth, "Neural nearest neighbors networks," *Advances in Neural information processing systems*, vol. 31, 2018.

[20] K. Zhang, W. Zuo, and L. Zhang, "Ffdnet: Toward a fast and flexible solution for cnn-based image denoising," *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4608–4622, 2018.

[21] X. Jia, S. Liu, X. Feng, and L. Zhang, "Focnet: A fractional optimal control network for image denoising," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 6054–6063.

[22] S. Anwar and N. Barnes, "Real image denoising with feature attention," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 3155–3164.

[23] D. Valsesia, G. Fracastoro, and E. Magli, "Deep graph-convolutional image denoising," *IEEE Transactions on Image Processing*, vol. 29, pp. 8226–8237, 2020.

[24] J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, and R. Timofte, "Swinir: Image restoration using swin transformer," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 1833–1844.

[25] K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Van Gool, and R. Timofte, "Plug-and-play image restoration with deep denoiser prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 10, pp. 6360–6376, 2021.

[26] Z. Liu, Y. Lin, Y. Cao, *et al.*, "Swin transformer: Hierarchical vision transformer using shifted windows," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 10 012–10 022.

[27] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Advances in neural information processing systems*, vol. 33, pp. 6840–6851, 2020.

[28] A. Hyvärinen and P. Dayan, "Estimation of non-normalized statistical models by score matching.," *Journal of Machine Learning Research*, vol. 6, no. 4, 2005.

[29] J. Lehtinen, J. Munkberg, J. Hasselgren, *et al.*, "Noise2noise: Learning image restoration without clean data," *arXiv preprint arXiv:1803.04189*, 2018.

[30] Y. Quan, M. Chen, T. Pang, and H. Ji, "Self2self with dropout: Learning self-supervised denoising from single image," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 1890–1898.

[31] Y. Xie, Z. Wang, and S. Ji, "Noise2same: Optimizing a self-supervised bound for image denoising," *Advances in neural information processing systems*, vol. 33, pp. 20 320–20 330, 2020.

[32] M. Papkov and P. Chizhov, "Swinia: Self-supervised blind-spot image denoising with zero convolutions," *arXiv preprint arXiv:2305.05651*, 2023.

[33] M. El Helou and S. Süsstrunk, "Blind universal bayesian image denoising with gaussian noise level learning," *IEEE Transactions on Image Processing*, vol. 29, pp. 4885–4897, 2020.

[34] S. Lefkimmiatis, "Universal denoising networks: A novel cnn architecture for image denoising," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 3204–3213.

[35] S. Gu, Y. Li, L. V. Gool, and R. Timofte, "Self-guided network for fast image denoising," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 2511–2520.

[36] X. Hou, H. Luo, J. Liu, *et al.*, "Learning deep image priors for blind image denoising," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1738–1747, 2019.

[37] H. Zhang, Y. Li, H. Chen, and C. Shen, "Memory-efficient hierarchical neural architecture search for image denoising," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3654–3663, 2019.

[38] C. Chen, Z. Xiong, X. Tian, Z.-J. Zha, and F. Wu, "Real-world image denoising with deep boosting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 12, pp. 3071–3087, 2019.

[39] L. Xu, J. Zhang, X. Cheng, F. Zhang, X. Wei, and J. Ren, "Efficient deep image denoising via class specific convolution," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, 2021, pp. 3039–3046.

[40] H. Zheng, H. Yong, and L. Zhang, "Deep convolutional dictionary learning for image denoising," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 630–641.

[41] S. Guo, Z. Yan, K. Zhang, W. Zuo, and L. Zhang, "Toward convolutional blind denoising of real photographs," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 1712–1722.

[42] H. Zhao, W. Shao, B. Bao, and H. Li, "A simple and robust deep convolutional approach to blind image denoising," *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, pp. 3943–3951, 2019.

[43] X. Wu, M. Liu, Y. Cao, D. Ren, and W. Zuo, "Unpaired learning of deep image denoising," in *European conference on computer vision*, Springer, 2020, pp. 352–368.

[44] J. W. Soh and N. I. Cho, "Deep universal blind image denoising," in *2020 25th International Conference on Pattern Recognition (ICPR)*, IEEE, 2021, pp. 747–754.

[45] S. Cheng, Y. Wang, H. Huang, D. Liu, H. Fan, and S. Liu, "Nbnet: Noise basis learning for image denoising with subspace projection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 4896–4906.

[46] M. Chang, Q. Li, H. Feng, and Z. Xu, "Spatial-adaptive network for single image denoising," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXX 16*, Springer, 2020, pp. 171–187.

[47] Y. Liu, Z. Qin, S. Anwar, *et al.*, "Invertible denoising network: A light solution for real noise removal," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 13 365–13 374.

[48] X. Hu, R. Ma, Z. Liu, *et al.*, "Pseudo 3d auto-correlation network for real image denoising," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 16 175–16 184.

[49] C. Ren, X. He, C. Wang, and Z. Zhao, "Adaptive consistency prior based deep network for image denoising," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 8596–8606.

[50] L. Chen, X. Lu, J. Zhang, X. Chu, and C. Chen, "Hinet: Half instance normalization network for image restoration," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 182–192.

[51] Z. Xia, M. Gharbi, F. Perazzi, K. Sunkavalli, and A. Chakrabarti, "Deep denoising of flash and no-flash pairs for photography in low-light environments," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 2063–2072.

[52] B. Moseley, V. Bickel, I. G. López-Francos, and L. Rana, "Extreme low-light environment-driven image denoising over permanently shadowed lunar regions with a physical noise model," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 6317–6327.

[53] X. Rui, X. Cao, Q. Xie, Z. Yue, Q. Zhao, and D. Meng, "Learning an explicit weighting scheme for adapting complex hsi noise," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 6739–6748.

[54] Z. Tu, H. Talebi, H. Zhang, *et al.*, "Maxim: Multi-axis mlp for image processing," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 5769–5780.

[55] B. Li, X. Liu, P. Hu, Z. Wu, J. Lv, and X. Peng, "All-in-one image restoration for unknown corruption," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 17 452–17 462.

[56] R. Ma, S. Li, B. Zhang, and Z. Li, "Generative adaptive convolutions for real-world noisy image denoising," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 36, 2022, pp. 1935–1943.

[57] H. Chen, Y. Wang, T. Guo, *et al.*, "Pre-trained image processing transformer," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 12 299–12 310.

[58] Z. Wang, X. Cun, J. Bao, W. Zhou, J. Liu, and H. U. Li, "A general u-shaped transformer for image restoration. arxiv 2021," *arXiv preprint arXiv:2106.03106*,

[59] J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, and R. Timofte, "Swinir: Image restoration using swin transformer," in *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, 2021, pp. 1833–1844. DOI: 10.1109/ICCVW54120.2021.00210.

[60] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, "Restormer: Efficient transformer for high-resolution image restoration," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 5718–5729. DOI: 10.1109/CVPR52688.2022.00564.

[61] X. Feng, H. Ji, W. Pei, J. Li, G. Lu, and D. Zhang, "U^2-former: Nested u-shaped transformer for image restoration via multi-view contrastive learning," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 34, no. 1, pp. 168–181, 2024. DOI: 10.1109/TCSVT.2023.3286405.

[62] Z. Deng, Y. Cai, L. Chen, *et al.*, "Rformer: Transformer-based generative adversarial network for real fundus image restoration on a new clinical benchmark," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 9, pp. 4645–4655, 2022. DOI: 10.1109/JBHI.2022.3187103.

[63] J. M. J. Valanarasu, R. Yasarla, and V. M. Patel, "Transweather: Transformer-based restoration of images degraded by adverse weather conditions," *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2343–2353, 2021.

[64] L. Liu, S. Yuan, J. Liu, X. Guo, Y. Yan, and Q. Tian, "Siamtrans: Zero-shot multi-frame image restoration with pre-trained siamese transformers," in *AAAI Conference on Artificial Intelligence*, 2021.

[65] J. Liang, J. Cao, Y. Fan, *et al.*, "Vrt: A video restoration transformer," *IEEE Transactions on Image Processing*, vol. 33, pp. 2171–2182, 2022.

[66] Z. Fu, Z. Fu, Z. Fang, *et al.*, "Prior skeleton based online deep reinforcement learning for coronary artery centerline extraction," *Proceedings of the Institution of Mechanical Engineers. Part H, Journal of Engineering in Medicine*, vol. 237, pp. 557–570, 2023.

[67] J. Lehtinen, J. Munkberg, J. Hasselgren, *et al.*, "Noise2noise: Learning image restoration without clean data," *ArXiv*, vol. abs/1803.04189, 2018.

[68] S. Cha and T. Moon, "Fully convolutional pixel adaptive image denoiser," *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 4159–4168, 2018.

[69] S. Cha and T. Moon, "Neural adaptive image denoiser," *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2981–2985, 2018.

[70] J. Liu, Y. Sun, X. Xu, and U. S. Kamilov, "Image restoration using total variation regularized deep image prior," *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 7715–7719, 2018.

[71] S. Cha, T. Park, B. Kim, J. Baek, and T. Moon, "Gan2gan: Generative noise learning for blind denoising with single noisy images," in *International Conference on Learning Representations*, 2019.

[72] J. D. Batson and L. A. Royer, "Noise2self: Blind denoising by self-supervision," *ArXiv*, vol. abs/1901.11365, 2019.

[73] S. Laine, T. Karras, J. Lehtinen, and T. Aila, "High-quality self-supervised deep image denoising," in *Neural Information Processing Systems*, 2019.

[74] M. Zhussip, S. Soltanayev, and S. Y. Chun, "Extending stein's unbiased risk estimator to train deep denoisers with correlated pairs of noisy images," in *Neural Information Processing Systems*, 2019.

[75] O. Sidorov and J. Y. Hardeberg, "Deep hyperspectral prior: Single-image denoising, inpainting, super-resolution," *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, pp. 3844–3851, 2019.

[76] J. Xu, Y. Huang, M.-M. Cheng, *et al.*, "Noisy-as-clean: Learning self-supervised denoising from corrupted image," *IEEE Transactions on Image Processing*, vol. 29, pp. 9316–9329, 2019.

[77] A. Krull, T. Viar, and F. Jug, "Probabilistic noise2void: Unsupervised content-aware denoising," in *Frontiers of Computer Science*, 2019.

[78] M. Prakash, M. Lalit, P. Tomanák, A. Krull, and F. Jug, "Fully unsupervised probabilistic noise2void," *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pp. 154–158, 2019.

[79] Y. Chen, C. Gao, E. Robb, and J.-B. Huang, "Nas-dip: Learning deep image prior with neural architecture search," *ArXiv*, vol. abs/2008.11713, 2020.

[80] N. Moran, D. Schmidt, Y. Zhong, and P. Coady, "Noisier2noise: Learning to denoise from unpaired noisy data," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12 061–12 069, 2019.

[81] Y. Xie, Z. Wang, and S. Ji, "Noise2same: Optimizing a self-supervised bound for image denoising," *ArXiv*, vol. abs/2010.11971, 2020.

[82] Y. Quan, M. Chen, T. Pang, and H. Ji, "Self2self with dropout: Learning self-supervised denoising from single image," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1887–1895, 2020.

[83] M. Papkov, K. Roberts, L. A. Madissoon, *et al.*, "Noise2stack: Improving image restoration by learning from volumetric data," in *MLMIR@MICCAI*, 2020.

[84] C. Niu, M. Li, F. Fan, *et al.*, "Suppression of correlated noise with similarity-based unsupervised deep learning," 2020.

[85] B. Yaman, S. A. H. Hosseini, and M. Akçakaya, "Noise2inpaint: Learning referenceless denoising by inpainting unrolling," *ArXiv*, vol. abs/2006.09450, 2020.

[86] K. Lee and W.-K. Jeong, "Noise2kernel: Adaptive self-supervised blind denoising using a dilated convolutional kernel architecture," *Sensors (Basel, Switzerland)*, vol. 22, 2020.

[87] T. Huang, S. Li, X. Jia, H. Lu, and J.-z. Liu, "Neighbor2neighbor: Self-supervised denoising from single noisy images," *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 14 776–14 785, 2021.

[88] T. Pang, H. Zheng, Y. Quan, and H. Ji, "Recorrupted-to-recorrupted: Unsupervised deep learning for image denoising," *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2043–2052, 2021.

[89] K. Kim and J.-C. Ye, "Noise2score: Tweedie's approach to self-supervised image denoising without clean images," in *Neural Information Processing Systems*, 2021.

[90] J. Lequyer, R. Philip, A. Sharma, W.-H. Hsu, and L. Pelletier, "A fast blind zero-shot denoiser," *Nature Machine Intelligence*, vol. 4, pp. 953–963, 2021.

[91] Y. Zeng, Y. Qian, Q. Zhang, J. Hou, Y. Yuan, and Y. He, "Idea-net: Dynamic 3d point cloud interpolation via deep embedding alignment," *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6328–6337, 2022.

[92] W. Lee, S. Son, and K. M. Lee, "Ap-bsn: Self-supervised denoising for real-world images via asymmetric pd and blind-spot network," *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 17 704–17 713, 2022.

[93] Z. Wang, J. Liu, G. Li, and H. Han, "Blind2unblind: Self-supervised image denoising with visible blind spots," *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2017–2026, 2022.

[94] C.-K. Ta, A. Aich, A. Gupta, and A. K. Roy-Chowdhury, "Poisson2sparse: Self-supervised poisson denoising from a single image," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2022.

[95] M. D. Chaity and M. A. N. I. Fahim, "Decay2distill: Leveraging spatial perturbation and regularization for self-supervised image denoising," *ArXiv*, vol. abs/2208.01948, 2022.

[96] X. Tian, Q. Wu, H. Wei, and Y. Zhang, "Noise2sr: Learning to denoise from super-resolved single noisy fluorescence image," *ArXiv*, vol. abs/2209.06411, 2022.

[97] E.-M. Hock, T.-O. Buchholz, A. Brachmann, F. Jug, and A. Freytag, "N2v2 - fixing noise2void checkerboard artifacts with modified sampling strategies and a tweaked network architecture," in *ECCV Workshops*, 2022.

[98] W. He, T. Uezato, and N. Yokoya, "Interpretable deep attention prior for image restoration and enhancement," *IEEE Transactions on Computational Imaging*, vol. 9, pp. 185–196, 2023.

[99] S. Chen, J. Zhang, Z. Yu, and T. Huang, "Exploring efficient asymmetric blind-spots for self-supervised denoising in real-world scenarios," 2023.

[100] Y. I. Jang, K.-O. Lee, G. Y. Park, S. Kim, and N. I. Cho, "Self-supervised image denoising with downsampled invariance loss and conditional blind-spot network," *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 12 162–12 171, 2023.

[101] D. Zhang, F. Zhou, Y. Jiang, and Z. Fu, "Mm-bsn: Self-supervised image denoising for real-world with multi-mask based on blind-spot network," *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 4189–4198, 2023.

[102] Y. Mansour and R. Heckel, "Zero-shot noise2noise: Efficient image denoising without any data," *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 14 018–14 027, 2023.

[103] M. Papkov and P. Chizhov, "Swinia: Self-supervised blind-spot image denoising with zero convolutions," *ArXiv*, vol. abs/2305.05651, 2023.

[104] R. Li, G. della Maggiora, V. Andriasyan, *et al.*, "Microscopy image reconstruction with physics-informed denoising diffusion probabilistic model," 2023.

[105] T.-A. Song, F. Yang, and J. Dutta, "Noise2void: Unsupervised denoising of pet images," *Physics in Medicine & Biology*, vol. 66, 2021.

[106] P. Reinagel and A. M. Zador, "Natural scene statistics at the centre of gaze.," *Network*, vol. 10 4, pp. 341–50, 1999.

[107] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing*, vol. 21, pp. 4695–4708, 2012.

[108] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a "completely blind" image quality analyzer," *IEEE Signal Processing Letters*, vol. 20, pp. 209–212, 2013.