

New method to improve the diagnostic utility of OCTA images in retinal disease

by

Rishav Bhardwaj

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Vision Science and Systems Design Engineering

Waterloo, Ontario, Canada, 2024

© Rishav Bhardwaj 2024

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

The following publications have resulted from the work presented in this thesis:

- Bhardwaj R, Jothi Balaji J, Lakshminarayanan V. OW-SLR: Overlapping Windows on Semi-Local Region for Image Super-Resolution. *J Imaging*. 2023 Nov 8;9(11):246. doi: 10.3390/jimaging9110246
- Bhardwaj R, Jothi Balaji J, Lakshminarayanan V. OW-SLR: overlapping windows on semi-local region for image super-resolution, *Proc. SPIE 12674, Applications of Digital Image Processing XLVI*, 1267416; <https://doi.org/10.1117/12.2680629>
- Bhardwaj R, Abdul Rasheed M, Jothi Balaji J, Lakshminarayanan V. Training method for the Removal of Motion Artifacts from OCTA images. In submission at *Journal of Medical Optics*.

Abstract

Purpose:

Diagnosing medical images necessitates years of experience to ensure accurate diagnoses. However, the current workforce available for this task falls significantly short compared to the volume of images requiring assessment. This places a considerable burden on the medical system during diagnosis. Additionally, medical images often contain artifacts, further complicating and prolonging the diagnostic process. This thesis serves as a solution to expedite diagnosis by enhancing the image quality of [Optical Coherence Tomography Angiography \(OCTA\)](#) images, thereby alleviating the strain on the system.

Aims:

1. **Method 1** (Chapter 2): Removal of motion artifacts from [OCTA](#) images. It is one of the toughest artifacts to be removed from an image.
2. **Method 2** (Chapter 3): Super-Resolution of [OCTA](#) image. Increasing the dimensions of the image and enhancing the quality to make diagnosis process efficient.

Conclusion: This work allows the removal of motion artifacts from the [OCTA](#) image and then enhance the quality of the image using super-resolution. In chapter 4 we show that the scatterplots were used to compare the correlations of the most commonly used parameters, [Foveal Avascular Zone \(FAZ\)](#) area, perimeter, and circularity index, between before and after super-resolution at $\times 2$ and $\times 3$ magnification. A p-value < 0.05 was considered significant for all statistical tests. Thus, making the diagnosis process simpler and better for medical practitioners.

Acknowledgements

I extend my heartfelt gratitude to the esteemed members of my thesis committee, Dr. Vasudevan Lakminarayanan, Dr. John Zelek, Dr. Kaamran Raahemifar, and Dr. Jeff Hovis. Their collective expertise, guidance, and commitment to excellence have been indispensable throughout the entire research process. My supervisors Dr. Vasudevan Lakminarayanan and Dr. John Zelek, deserve special acknowledgment for their unwavering support and scholarly insights. Their mentorship has been instrumental in shaping the direction and focus of my research. Dr. Lakminarayanan's expertise in Optics and Dr. Zelek's dedication to pushing intellectual boundaries have profoundly influenced the quality and depth of this work.

I would like to express my gratitude to Mr. Jothi Balaji, whose pivotal contribution significantly influenced the trajectory of this thesis. Mr. Balaji played a crucial role in data collection and utilized his expertise in clinical skills to perform segmentation of motion artifacts on the gathered data. I am sincerely grateful to Dr. Kaamran Raahemifar and Dr. Jeff Hovis for their valuable contributions to my thesis. Their thoughtful feedback, rigorous examination of my research, and constructive criticism have played a pivotal role in refining the methodology and conclusions of this study.

I would also like to express my appreciation to TEEL Lab, for fostering a stimulating academic environment. The collaborative spirit and intellectual exchange within the department have significantly enriched my understanding of the computer vision field. My heartfelt gratitude is owed to Dr. Abdul Rasheed Mohammed, for ensuring the timely completion of motion artifact annotations, without which I would not have been able to finish my thesis on schedule.

I am deeply indebted to my family for their steadfast encouragement and unwavering support throughout my academic journey. Despite facing financial challenges, they prioritized my education, laying the groundwork for my academic successes. I am very grateful for their unwavering belief in me, and I take great pride in the fact that I have lived up to their expectations.

Finally, I thank all my wonderful friends from Iran without whom this entire experience would never have been this amazing. They have not only taught me invaluable lessons about life but have also helped me grow into a better person compared to who I was 16 months ago. From learning Persian to engaging in heartfelt discussions during our Bishinim (Persian: sitting together), from sharing cultural experiences and enjoying delicious food to hitting the gym together, spending almost every moment of the day with them has made me feel like I am almost half-Iranian myself. Their friendship has been an incredible blessing, and I am grateful for the impact they have had on my life.

Dedication

To my loving parents, my hardworking elder brother and my cat, Mini.

Table of Contents

Author's Declaration	ii
Statement of Contributions	iii
Abstract	iv
Acknowledgements	v
Dedication	vi
List of Figures	x
List of Tables	xii
List of Abbreviations	xiii
1 Introduction	1
1.1 Retina	2
1.1.1 Structure of Retina	2
1.1.2 Functioning of Retina	3
1.2 Optical Coherence Tomography	6
1.2.1 Basic Principle	6
1.2.2 Types of OCT	6

1.2.3	Applications of OCT	10
1.3	Optical Coherence Tomography Angiography	11
1.3.1	Principles of OCTA	11
1.3.2	Applications of OCTA	12
1.4	Different types of Artifacts in OCTA images	12
1.4.1	Projection Artifacts	12
1.4.2	Shadow Artifacts	13
1.4.3	Bulk Motion Artifact	14
1.4.4	Motion Artifact	14
2	Removal of Motion Artifacts	16
2.1	Overview	17
2.2	Introduction	17
2.3	Method	18
2.3.1	Data Collection and Annotation	18
2.3.2	Image Registration	19
2.3.3	Generating faux motion artifacts	20
2.3.4	Cardinal mask	21
2.3.5	Working	22
2.3.6	Loss Function	22
2.4	Experiments	22
2.4.1	Dataset	22
2.4.2	Setup	22
2.4.3	Quantitative Results	23
2.4.4	Qualitative Results	23
2.5	Conclusion	23

3	Super-Resolution of the OCTA image	25
3.1	Overview	26
3.2	Introduction	26
3.3	Related Work	27
3.4	Method	29
3.4.1	Backbone Framework	29
3.4.2	Locating the Semi-Local Region	30
3.4.3	Overlapping Windows	31
3.5	Results and Discussion	33
3.5.1	Dataset	33
3.5.2	Implementation Details	33
3.5.3	Quantitative Results	34
3.6	Conclusion	36
4	Evaluation of Clinical Parameters	38
4.1	Introduction	39
4.2	Software	39
4.3	Data	40
4.4	Experiment	40
4.5	Result and Discussion	40
5	Conclusion and future research	44
	References	46

List of Figures

1.1	Layers in Retina. Image taken from [11] is licensed under CC BY.	5
1.2	A Typical OCT Image with Retinal Layers.	6
1.3	The Schematic diagrams of A:TD-OCT, B:SD-OCT, C: SS-OCT and D:FF-OCT machines. Image taken from [69].	10
2.1	(a) and (b) are two images taken of the same eye at different timestamps. (c) and (d) consist of the annotation of motion artifacts in (a) and (b) respectively by the optometrist.	19
2.2	(a) is now aligned with (b) using homography[12]	20
2.3	(a) and (b) are two images cropped from the Fig 2.2. (c) shows how a thin strip is taken from (b) and pasted in (a), replicating a motion artifact. (c) consisting of motion artifact is served as one of the inputs to the model and (a) acts as a ground truth image with no artifact.	21
2.4	(a) shows the result after using Otsu’s Thresholding [41]. (b) is the mask highlighting the region where the motion artifact exists. (c) is produced by multiplying (a) with (b).	21
2.5	Red arrow shows the part of the image consisting of motion artifact. Green arrow shows how the architecture removes the motion artifact.	24
3.1	(a) An high resolution (LR) image is taken. (b) It is passed through EDSR [37] and a feature map is produced. (c) Locating the semi-local region ($M = 6$) around a random selected point from high resolution (HR) image. (d) Semi-local region is passed through the proposed Overlapping Windows. (e) This output is passed through the Multilayer Perceptron (MLP) to give out the Red, Green and Blue (RGB) value of a randomly selected point. Steps (c-e) are performed for all the points in the HR image.	30

3.2	To extract features from a feature map of size 3×3 , we focus on a specific query point represented by a red dot. In order to determine which pixel locations in the feature map correspond to this query point, we compute the Euclidean distance between the query point and the center points of each pixel location. In the provided image, the black line represents the closest pixel location in the feature map to the query point.	32
3.3	(a) Given an HR image, a point of interest (red dot) is selected to predict its RGB value. (b) Its corresponding spatially equivalent 2D coordinate is selected from the feature map. (c) Locating the semi-local region ($M = 6$) around the calculated 2D coordinate.	33
3.4	The first iteration of overlapping windows, where the window size = $M - 1$ ($M = 6$). Assuming the feature map is of negligible depth and four windows are positioned at the four corners of the feature map.	34
3.5	A 96×96 patch is taken and its size is reduced to 24×24 (first row), 32×32 (second row) and 48×48 (third row) using bicubic interpolation. Our architecture uses the same set to weights reproduce the given results. However, others require different set of weights for a newer scale to be trained on. The Peak signal-to-noise ratio (PSNR) results of each image are shown in Table 3.1.	35
4.1	FAZSeg[60] software being used to find the 15 parameters.	39
4.2	Correlation between FAZ area measured before and after super-resolution.	42
4.3	Correlation between FAZ perimeter measured before and after super-resolution.	42
4.4	Correlation between FAZ circularity index measured before and after super-resolution.	43

List of Tables

2.1	PSNR and Structural Similarity (SSIM) results across various architectures on the proposed training method.	23
2.2	Training time and different architectures.	23
3.1	PSNR result of each of the input images across different methods shown in Figure 3.5.	36
3.2	Time taken to extrapolate a 320×320 image on a single Nvidia Titan V of 12 Gigabyte size.	36
3.3	PSNR result on the 300 images from OCT500 [35]	37
4.1	A comparison of FAZ parameters before and after super-resolution.	41

List of Abbreviations

- CNNs** Convolutional Neural Networks [28](#)
- EDSR** Enhanced Deep Residual Networks [29](#), [34](#)
- ESRGAN** Enhanced Super-Resolution Generative Adversarial Network [29](#)
- FAZ** Foveal Avascular Zone [iv](#), [xii](#), [39–41](#)
- FD-OCT** Frequency-Domain OCT [7](#)
- FF-OCT** Full-Field OCT [8](#)
- GAN** Generative Adversarial Networks [28](#)
- GPU** Graphics processing unit [22](#)
- GT** Ground Truth [35](#)
- HR** high resolution [x](#), [xi](#), [26](#), [27](#), [29–31](#), [33](#), [36](#), [44](#)
- INL** Inner Nuclear Layer [2](#), [4](#)
- IPL** Inner Plexiform Layer [3](#)
- IVOCT** Intravascular OCT [11](#)
- LIIF** Local Implicit Image Function [26](#), [36](#)
- LR** high resolution [x](#), [26–30](#), [34](#), [35](#), [44](#)

MLP Multilayer Perceptron [x](#), [30](#), [32](#), [37](#), [44](#)

OCE Optical Coherence Elastography [8](#)

OCT Optical Coherence Tomography [6](#), [7](#), [10](#), [11](#), [17](#)

OCTA Optical Coherence Tomography Angiography [iv](#), [viii](#), [11–15](#), [17](#), [18](#), [22–24](#), [26](#), [33](#), [36](#), [39](#), [44](#)

ONL Outer Nuclear Layer [2](#)

OPL Outer Plexiform Layer [2](#)

OW-SLR Overlapping Windows on Semi-Local Region [26](#), [34](#)

PS-OCT Polarization-Sensitive OCT [8](#)

PSNR Peak signal-to-noise ratio [xi](#), [xii](#), [23](#), [33](#), [35–37](#)

RGB Red, Green and Blue [x](#), [26](#), [27](#), [29–31](#), [33](#), [37](#), [44](#)

RPE Retinal Pigment Epithelium [2](#), [13](#)

SD-OCT Spectral-Domain OCT [9](#)

SR super-resolution [26–28](#)

SS-OCT Swept-Source OCT [7](#)

SSIM Structural Similarity [xii](#), [23](#)

TD-OCT Time-Domain OCT [6](#), [7](#)

VAD Vessel Avascular Density [41](#)

VDI Vessel Diameter Index [41](#)

Chapter 1

Introduction

1.1 Retina

The retina, a delicate and intricate layer of tissue lining the inner surface of the eye, plays a pivotal role in the process of vision. Comprising several layers with specialized functions, the retina is essential for converting light stimuli into neural signals that the brain can interpret as vision. This article delves into the structure and functioning of the retina, exploring the layers that contribute to its remarkable capabilities.

1.1.1 Structure of Retina

The retina can be likened to the film in a camera, capturing visual information and transmitting it to the brain for interpretation. Structurally, the retina is composed of several layers, each with distinct roles in the visual process.

1. **Retinal Pigment Epithelium (RPE):**

At the outermost layer of the retina lies the **RPE**. This layer is responsible for nourishing the photoreceptor cells, maintaining the health of the outer segments of photoreceptors, and absorbing excess light to prevent scattering within the eye.

2. **Photoreceptor Layer :**

The photoreceptor layer contains two main types of cells - rods and cones. Rods are sensitive to low light conditions and facilitate peripheral vision, while cones operate in bright light and are responsible for color vision. Photoreceptor cells capture light and convert it into electrical signals.

3. **Outer Nuclear Layer (ONL):**

The **ONL** contains the nuclei of the photoreceptor cells, playing a crucial role in the production and maintenance of these light-sensitive cells.

4. **Outer Plexiform Layer (OPL):**

The **OPL** is where synapses occur between photoreceptor cells and bipolar cells and horizontal cells are present. It serves as a crucial site for signal processing before transmitting information to the inner layers of the retina.

5. **Inner Nuclear Layer (INL):**

The **INL** consists of various cell bodies of interneurons, including bipolar cells, horizontal cells, and amacrine cells. These cells play vital roles in modulating and processing visual information before it reaches the ganglion cells.

6. **Inner Plexiform Layer (IPL):**

The IPL is where synapses take place between bipolar cells, amacrine cells, and ganglion cells. This layer is critical for refining visual information and transmitting it to the ganglion cells.

7. **Ganglion Cell Layer:**

Ganglion cells are the final output neurons of the retina, sending processed visual information to the brain through the optic nerve. Some ganglion cells are specialized for detecting motion, while others contribute to the perception of color and contrast.

8. **Nerve Fiber Layer:**

The nerve fiber layer contains the axons of ganglion cells, which bundle together to form the optic nerve. This layer is the conduit through which visual information travels from the retina to the brain.

1.1.2 Functioning of Retina

Phototransduction in Photoreceptor Cells:

At the heart of retinal function lies phototransduction, the process by which light signals are converted into electrical signals. This occurs primarily in the outer segments of the photoreceptor cells – rods and cones. The outer segments contain stacks of membranous disks that house photopigments, proteins sensitive to light. When struck by photons, these pigments undergo structural changes, initiating a cascade of events that culminate in the generation of electrical signals.

In rods, the photopigment rhodopsin plays a central role, while cones possess different photopigments responsible for color vision – short-wavelength cones for blue, middle-wavelength cones for green, and long-wavelength cones for red. The variations in these pigments contribute to the perception of a diverse spectrum of colors.

Signal Processing in the Inner Layers:

Once the photoreceptor cells capture light and generate electrical signals, the information undergoes intricate processing in the inner layers of the retina. Bipolar cells, positioned in the inner nuclear layer, receive signals from photoreceptor cells and serve as intermediaries in transmitting information to the ganglion cells.

Horizontal cells, found in the [INL](#), modulate signals laterally, allowing for the enhancement of contrasts and the sharpening of edges in the visual scene. Amacrine cells, present in both the inner nuclear layer and inner plexiform layer, contribute to the processing of motion and other complex visual features.

Synaptic Transmission in the Inner Plexiform Layer:

The inner plexiform layer is a crucial site for synaptic transmission, where bipolar cells synapse with both amacrine cells and ganglion cells. This synaptic network allows for the integration of signals and the extraction of essential features from the visual input. Amacrine cells play a pivotal role in shaping the receptive fields of ganglion cells, influencing their responses to different aspects of visual stimuli.

Ganglion Cells: The Conduits to the Brain:

Ganglion cells, situated in the ganglion cell layer, are the final output neurons of the retina. These cells have diverse functional properties, with some responding to changes in luminance, others specializing in detecting motion, and yet others contributing to the perception of color and form. The axons of ganglion cells converge to form the optic nerve, which carries the compiled visual information to the brain for further processing.

Retinal Processing and Visual Perception:

The complex interactions within the retinal circuitry result in the extraction of fundamental features from the visual input. The processing of visual information includes edge detection, contrast enhancement, and motion perception. The parallel processing of different aspects of visual stimuli allows for the integration of diverse features, contributing to our holistic perception of the visual scene.

Adaptation and Dynamic Range:

The retina exhibits adaptive mechanisms to accommodate a wide range of light intensities. Photoreceptor cells adjust their sensitivity to light levels through processes like photopigment regeneration and changes in the responsiveness of synaptic connections. This adaptation ensures that the retina can function optimally in both bright and dim lighting conditions, providing us with a consistent perception of the visual environment.

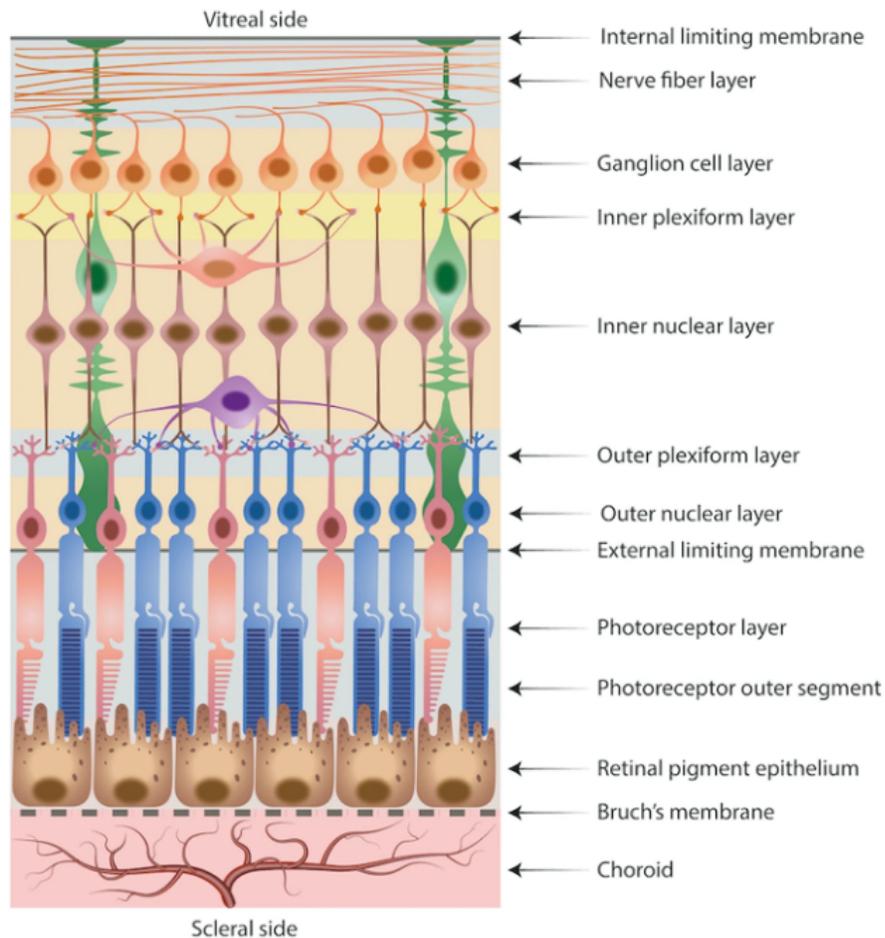


Figure 1.1: Layers in Retina. Image taken from [11] is licensed under [CC BY](#).

In conclusion, the functioning of the retina is a finely orchestrated symphony of cellular and molecular processes. From the initial capture of light by photoreceptor cells to the complex processing in the inner layers and the transmission of refined signals through ganglion cells, the retina demonstrates an incredible capacity for visual information processing. Understanding these intricacies not only deepens our appreciation for the physiological marvel of vision but also offers insights into the remarkable adaptability and efficiency of biological systems. Figure 1.1 shows the pictorial representation of different layers that exist in retina.

1.2 Optical Coherence Tomography

Optical Coherence Tomography (OCT) is a non-invasive imaging technique that uses light waves to capture detailed, cross-sectional images of biological tissues. It was first developed in the early 1990s and has since become a valuable tool in various medical fields, particularly ophthalmology and cardiology. OCT enables high-resolution, real-time imaging of biological tissues at micrometer-scale resolution.

1.2.1 Basic Principle

OCT is based on the principle of interferometry, specifically low-coherence interferometry. The key components include a light source, a beamsplitter, a reference arm, and a sample arm. A beam of light is split into a reference arm and a sample arm. The light reflected from the tissue and the reference arm is recombined, and interference patterns are analyzed to generate detailed images.

Figure 1.2 is an OCT image of the different retinal layers.

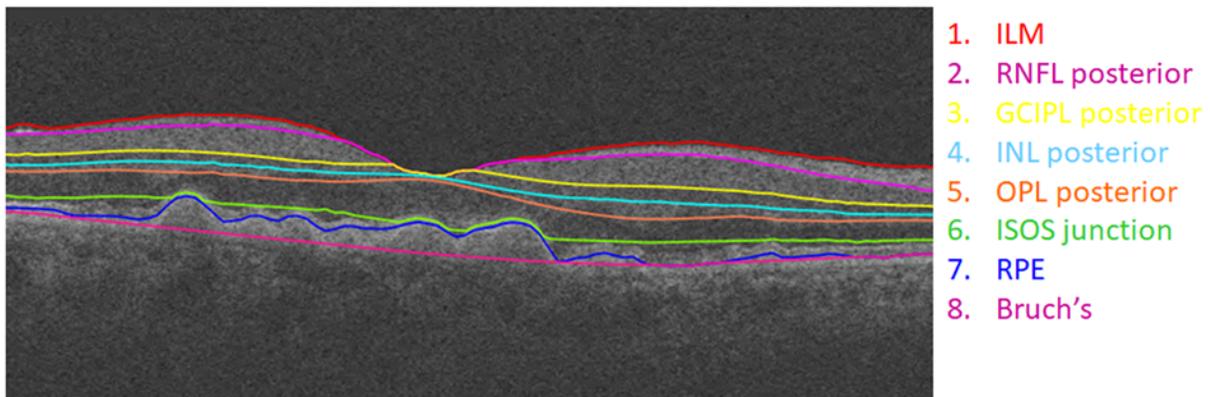


Figure 1.2: A Typical OCT Image with Retinal Layers.

1.2.2 Types of OCT

Different types of OCT are discussed here.

1. **Time-Domain OCT (TD-OCT)**[5]:

- **Principle:** [TD-OCT](#) is the original form of [OCT](#) and operates by physically scanning the reference arm's mirror to vary the optical path length.
- **Advantages:**
 - Proven technology with a long history of use.
 - Suitable for various applications, especially in ophthalmology.
- **Limitations:**
 - Limited imaging speed due to mechanical scanning.
 - Susceptible to motion artifacts.

2. [Frequency-Domain OCT \(FD-OCT\)](#)^[7]:

- **Principle:** [FD-OCT](#) eliminates the need for mechanical scanning by using a spectrometer to measure the interference spectrum. It is also known as Fourier-Domain OCT.
- **Types of FD-OCT:**
 - **Spatially Encoded Frequency-Domain OCT** ^[3]: Utilizes spatial encoding techniques to improve the imaging speed and sensitivity.
- **Advantages:**
 - Higher imaging speed compared to [TD-OCT](#).
 - Improved sensitivity and signal-to-noise ratio.
- **Limitations:**
 - More complex instrumentation.

3. [Swept-Source OCT \(SS-OCT\)](#)^[24]:

- **Principle:** [SS-OCT](#) employs a tunable laser source that rapidly sweeps through a range of wavelengths. It is a specific type of [FD-OCT](#).
- **Advantages:**
 - Improved imaging depth and speed.
 - Reduced sensitivity roll-off with depth.
 - Suitable for imaging highly scattering tissues.
- **Limitations:**
 - Requires stable, high-speed tunable lasers.

- Susceptible to motion artifacts.

4. **Full-Field OCT (FF-OCT)**[26]:

- **Principle:** In **FF-OCT**, the entire field of view is illuminated simultaneously, and the interference pattern is analyzed for each point in the field.
- **Advantages:**
 - High-speed imaging over a large field of view.
 - Suitable for imaging biological tissues in vivo.
- **Limitations:**
 - Complex instrumentation.
 - Limited depth penetration.

5. **Polarization-Sensitive OCT (PS-OCT)**[6, 4]:

- **Principle:** **PS-OCT** measures the polarization states of the backscattered light to provide additional contrast information.
- **Applications:**
 - Differentiation of tissues with varying birefringence (e.g., muscle and nerve fibers).
 - Detection of pathological changes in tissue structure.
- **Advantages:**
 - Enhanced tissue contrast.
- **Limitations:**
 - Requires specialized components.

6. **Optical Coherence Elastography (OCE)**[38]:

- **Principle:** Combines OCT with elastography to assess tissue mechanical properties.
- **Applications:**
 - Detection of tissue stiffness changes, useful in cancer diagnosis.
 - Monitoring changes in tissue elasticity during interventions.
- **Advantages:**

- Provides additional functional information.

- **Limitations:**

- Limited by the need for specialized elastography techniques.

7. Multi-Modal OCT[31]:

- **Principle:** Combines OCT with other imaging modalities (e.g., fluorescence imaging, confocal microscopy) to obtain complementary information.

- **Applications:**

- Enhanced tissue characterization.
- Improved identification of specific structures or molecules.

- **Advantages:**

- Comprehensive information.

- **Limitations:**

- Increased complexity.

8. Spectral-Domain OCT (SD-OCT)[30, 20]:

- **Principle:** In SD-OCT, a broadband light source is used to illuminate the tissue, and the backscattered light is interfered with a reference beam. Instead of physically scanning the reference arm, a spectrometer is employed to analyze the interference spectrum. The spectrum is then transformed using a Fourier transform, revealing depth information about the sample.

- **Applications:**

- Enhanced tissue characterization.
- Improved identification of specific structures or molecules.

- **Advantages:**

- High Imaging Speed.
- Improved Sensitivity.
- Reduced Motion Artifacts.
- Increased Depth Resolution.
- Real-Time Imaging.
- Versatility.

- **Limitations:**

- Instrumentation can be complex.
- There may be trade-offs between imaging depth, resolution, and speed.
- The need for stable light sources and calibration is crucial for obtaining accurate and reliable results.

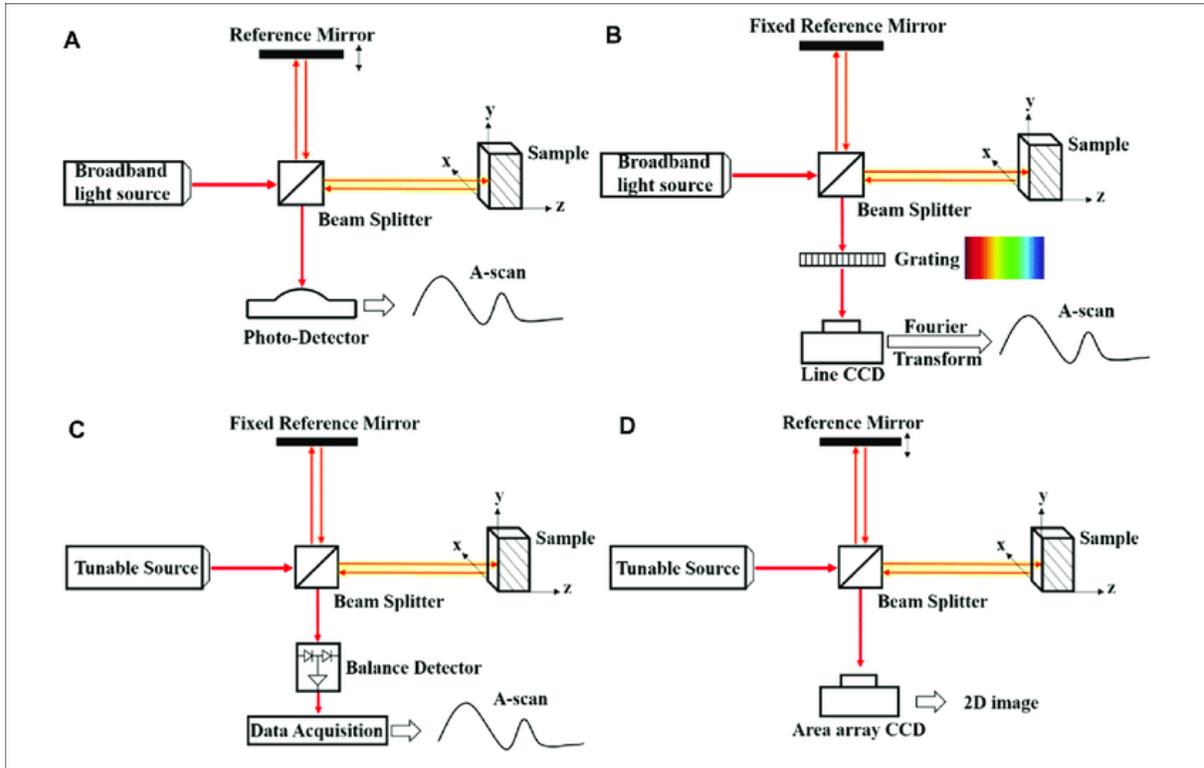


Figure 1.3: The Schematic diagrams of A:TD-OCT, B:SD-OCT, C: SS-OCT and D:FF-OCT machines. Image taken from [69].

1.2.3 Applications of OCT

1. Ophthalmology:

- **Retinal Imaging:** OCT is extensively used for imaging the retina, enabling visualization of the different layers of the retina and aiding in the diagnosis and management of various eye diseases such as macular degeneration, diabetic retinopathy, and glaucoma.

- **Corneal Imaging:** OCT is employed for imaging the cornea, providing detailed information about corneal thickness and morphology.

2. Cardiology:

- **Intravascular OCT (IVOCT):** This application involves imaging blood vessels from within, providing high-resolution images of arterial structures. It is particularly useful in assessing coronary artery disease and guiding interventions.
- **Corneal Imaging:** OCT is employed for imaging the cornea, providing detailed information about corneal thickness and morphology.

3. Other Medical Fields:

- **Dermatology:** OCT is used for skin imaging, assisting in the diagnosis of skin conditions and monitoring treatment responses.
- **Gastroenterology:** OCT can be used for imaging the gastrointestinal tract, aiding in the detection and characterization of lesions.

1.3 Optical Coherence Tomography Angiography

OCTA is an advanced imaging technique that provides high-resolution, non-invasive visualization of the vasculature in various tissues. It has revolutionized the field of ophthalmology, providing detailed images of retinal and choroidal blood flow without the need for contrast dyes or invasive procedures. OCTA utilizes principles of traditional OCT combined with angiography to produce volumetric images of blood vessels.

1.3.1 Principles of OCTA

1. **Interference-Based Imaging:** Similar to conventional OCT, OCTA relies on low-coherence interferometry to measure the backscattered light from tissue. By comparing the interference patterns between consecutive scans, it can detect motion caused by blood flow.
2. **Blood Flow Detection:** OCTA distinguishes static tissue from flowing blood by detecting variations in signal intensity over time. This enables the visualization of both structural and blood flow information.

3. **Depth-Resolved Imaging:** With advancements in technology, [OCTA](#) can provide depth-resolved images, allowing visualization of blood flow at different layers of the tissue. This capability is particularly useful in assessing various retinal and choroidal pathologies.

1.3.2 Applications of [OCTA](#)

1. **Retinal Diseases:** [OCTA](#) has become an invaluable tool in the diagnosis and management of retinal diseases such as diabetic retinopathy, age-related macular degeneration, and retinal vascular occlusions. It enables clinicians to visualize microvascular changes associated with these conditions, facilitating early detection and treatment monitoring.
2. **Glaucoma:** [OCTA](#) aids in assessing the optic nerve head and peripapillary vasculature, providing insights into the pathophysiology of glaucoma. It helps in identifying vascular abnormalities associated with glaucomatous optic neuropathy and monitoring disease progression.
3. **Choroidal Imaging:** [OCTA](#) allows visualization of the choroidal vasculature, aiding in the evaluation of choroidal neovascularization in age-related macular degeneration and other choroidal diseases. It provides detailed information about the location, extent, and flow characteristics of these abnormal vessels, guiding treatment decisions.

1.4 Different types of Artifacts in [OCTA](#) images

[OCTA](#) images capture blood flow in the retina and are utilized in ophthalmology for diagnosing various eye conditions. Different types of artifacts can occur in [OCTA](#) images, which can affect the interpretation and analysis of the images. Here are some common types of artifacts found in [OCTA](#) images:

1.4.1 Projection Artifacts

Projection artifacts, also known as decorrelation tail artifacts, are a common occurrence in [OCTA](#) imaging. These artifacts arise from the projection of blood vessels from superficial retinal layers onto deeper layers, leading to false positive signals and potentially misleading interpretations of the vasculature.

The phenomenon of projection artifacts has been extensively studied in the field of ophthalmology. A study by [56] investigated the characteristics of projection artifacts in OCTA images and proposed methods for their identification and mitigation. The authors emphasized the importance of distinguishing between true flow signals and artifacts to avoid misinterpretation of retinal vascular patterns.

Moreover, the study conducted by [65] delved into the development of advanced algorithms for the suppression of projection artifacts in OCTA images. By employing complex signal processing techniques and machine learning algorithms, they demonstrated significant improvements in artifact reduction, thereby enhancing the reliability of OCTA imaging for clinical diagnosis and research purposes.

Projection artifacts pose a challenge in the accurate interpretation of OCTA images, particularly in assessing the vascular morphology and detecting pathological changes in the retina. Clinicians and researchers need to be aware of these artifacts and utilize appropriate strategies to minimize their impact on image analysis and interpretation.

1.4.2 Shadow Artifacts

OCTA images occur due to reduced signal penetration through retinal layers, often caused by dense retinal pigment or blood vessels casting shadows onto deeper layers. These artifacts can obscure underlying structures and affect the interpretation of vasculature patterns in the retina.

Research in the field of ophthalmology has shed light on the characteristics and implications of shadow artifacts in OCTA imaging. For instance, the study by [19] highlighted the impact of shadowing from overlying retinal vessels on the visualization of deeper retinal layers in OCTA images. The authors emphasized the need for strategies to mitigate shadow artifacts to improve the accuracy of retinal vascular analysis.

Furthermore, the work of [21] investigated the influence of RPE alterations on shadow artifacts in OCTA images. They observed that disruptions in the RPE, such as atrophy or hypertrophy, can exacerbate shadowing effects, leading to decreased image quality and potential misinterpretation of retinal vasculature.

To address the challenge posed by shadow artifacts, advanced image processing techniques have been developed. For instance, the study by [66] proposed a method for shadow artifact removal in OCTA images based on adaptive thresholding and morphological operations. Their findings demonstrated the efficacy of this approach in enhancing image clarity and facilitating accurate vascular analysis.

In summary, shadow artifacts represent a significant challenge in OCTA imaging, impacting the visualization and interpretation of retinal vasculature. Recognizing these artifacts and employing appropriate correction methods are essential for improving the reliability of OCTA-based diagnosis and research.

1.4.3 Bulk Motion Artifact

Bulk motion artifacts are a common issue in OCTA imaging, arising from sudden shifts or movements of the entire eye during image acquisition. These artifacts can distort the image and lead to inaccuracies in the assessment of retinal vasculature.

Several studies have investigated the characteristics and implications of bulk motion artifacts in OCTA imaging. For example, the study by Gao et al. [15] highlighted the impact of eye movement on OCTA image quality and the occurrence of bulk motion artifacts, particularly in cases where patients are unable to maintain stable fixation during image acquisition. The authors emphasized the need for effective motion correction techniques to mitigate the influence of bulk motion artifacts on OCTA image interpretation.

Moreover, the study by [9] focused on the development of algorithms for real-time motion correction in OCTA imaging. By implementing sophisticated image processing techniques and motion tracking algorithms, they demonstrated significant improvements in artifact reduction and image quality enhancement.

Additionally, research efforts have been directed towards the integration of eye-tracking systems into OCTA devices to minimize bulk motion artifacts. The study by [66] investigated the efficacy of an integrated eye-tracking system for real-time motion correction during OCTA image acquisition. Their findings highlighted the potential of eye-tracking technology in reducing bulk motion artifacts and improving the accuracy of retinal vascular analysis.

Bulk motion artifacts pose a significant challenge in OCTA imaging, impacting the quality and reliability of vascular assessment. Recognizing these artifacts and implementing appropriate correction methods are essential for ensuring accurate diagnosis and research outcomes in ophthalmology.

1.4.4 Motion Artifact

Motion artifacts represent a significant challenge in OCTA imaging, resulting from involuntary eye movements during image acquisition. These artifacts can lead to blurred or distorted images, affecting the interpretation of retinal vasculature.

Several studies have investigated motion artifacts in OCTA imaging and proposed methods for their mitigation. For instance, the study by [56] examined the impact of motion artifacts on OCTA image quality and emphasized the importance of motion tracking and correction techniques to improve image clarity and reliability. They highlighted the need for real-time monitoring of eye movements to minimize motion artifacts during image acquisition.

Furthermore, the work by [49] focused on the development of algorithms for motion artifact detection and correction in OCTA images. By utilizing advanced image processing techniques and motion tracking algorithms, they demonstrated significant improvements in artifact reduction and image quality enhancement.

Moreover, the study by [14] investigated the efficacy of deep learning-based approaches for motion artifact removal in OCTA images. By training convolutional neural networks on a large dataset of OCTA images with motion artifacts, they achieved promising results in automatically detecting and correcting motion artifacts, thereby improving image quality and diagnostic accuracy.

Motion artifacts pose a significant challenge in OCTA imaging, impacting the reliability and accuracy of vascular analysis. Recognizing these artifacts and implementing appropriate correction methods are crucial for ensuring the quality of OCTA-based diagnosis and research in ophthalmology.

Chapter 2

Removal of Motion Artifacts

This chapter is based on:

1. Bhardwaj R, Abdul Rasheed M, Jothi Balaji J, Lakshminarayanan V. Training method for the Removal of Motion Artifacts from OCTA images. In submission at Journal of Medical Optics.

2.1 Overview

There are different kinds of artifacts found in OCTA images which make the entire diagnosis process of retinal diseases by the medical practitioners complicated. In this work, we propose a training technique which is able to simply as well as efficiently remove motion artifacts from the OCTA images. We generate artificial motion artifacts in an image which helps us to create a training image pair of one with and without motion artifacts in it. We also show the generation of cardinal mask which plays a pivotal role in the removal of these artifacts. This training method is used with other architectures which makes it a plug-and-play system in the entire removal process.

2.2 Introduction

OCT is a rapid and non-invasive optical imaging technology utilized not only in laboratory research but also in clinical applications [18]. Its extension, OCTA, has found application in diverse fields such as ophthalmology, dermatology, and brain functional imaging. For diagnosing retinal conditions like diabetic retinopathy [52], macular degeneration[61], and choroidal neovascular membrane[46], OCTA plays a crucial role. While OCTA plays a pivotal role in the mentioned applications, it is highly susceptible to various types of artifacts. These artifacts result in the loss of crucial information from the images, complicating the diagnosis process for medical practitioners. Numerous approaches [13, 28, 32] have been introduced to address various types of artifacts; however, mitigating motion artifacts in OCTA images remains one of the most challenging tasks.

Motion artifacts in OCTA images pose significant challenges to the diagnosis process for medical practitioners. These artifacts can arise from involuntary eye movements, patient motion, or even instability in the imaging system itself. They manifest as blurring, distortion, or ghosting in the acquired images, leading to the loss of critical structural and vascular details. This loss of information impedes accurate interpretation and diagnosis of retinal conditions such as diabetic retinopathy, macular degeneration, and choroidal neovascular membrane. For instance, a study by [14] demonstrated that motion artifacts in OCTA images can obscure subtle changes in retinal vasculature, potentially masking the presence of pathological features. Similarly, the work by [56] highlighted how motion artifacts can affect the assessment of choroidal vasculature in diseases like age-related macular degeneration. Consequently, addressing motion artifacts in OCTA imaging is essential for improving diagnostic accuracy and enhancing clinical outcomes.

In this paper, we try to tackle the problem of motion artifacts in the OCTA images. We propose a training process that can be used with architectures like UNet[48], UNet++[70] and ENet[43] to remove motion artifacts. In this proposed training method, we produce artificial motion artifacts in a OCTA images that remarkably replicate real motion artifacts found in an OCTA image. Using this technique we are able to achieve a training pair, with and without the artifact, which makes the entire feature learning process for the architecture simpler. First, we collect two images of the same eye of a patient taken at different timestamps. Both the images would never perfectly align and a disparity would exist between the two images due to the natural motion of the eye. Upon aligning, we horizontally split the image into multiple images each with smaller heights. An image pair is selected from the same region of the two aligned images. Then, we crop a thin stripe of image from one of the pair and paste it into the other at the same aligned location. We also generate a cardinal mask along the stripe which takes the non artifact region of the image into account. This cardinal mask plays a very crucial role in the removal of motion artifact. The synthetically generated motion artifact image along with the cardinal mask serves as an input to the network. This simple yet efficient technique is able to achieve promising results in removing motion artifacts from the OCTA image.

2.3 Method

In this section, we delineate the steps involved for the removal of motion artifacts from the OCTA image. In 2.3.1, we elucidate the data collection steps and how it is annotated. In 2.3.2, the two images taken of the same eye are aligned using image registration. In 2.3.3, we show the main part of this work in which we generate artificial motion artifacts in OCTA images for training. Finally, in 2.3.4, we explain the creation of cardinal mask which plays a crucial role in the removal of motion artifacts from the image.

2.3.1 Data Collection and Annotation

In this step, we take two images of the same eye of a patient taken at two different timesteps. Since the eye undergoes natural motion, there would exist some kind of disparity between the two images. Then, the regions having the motion artifacts in the collected images are annotated by the optometrists. This annotation will also be used in generating the cardinal mask, explained in 2.3.4, which will play a pivotal role for the removal of artifact. Pictorial representation of this step is shown in Fig 2.1.

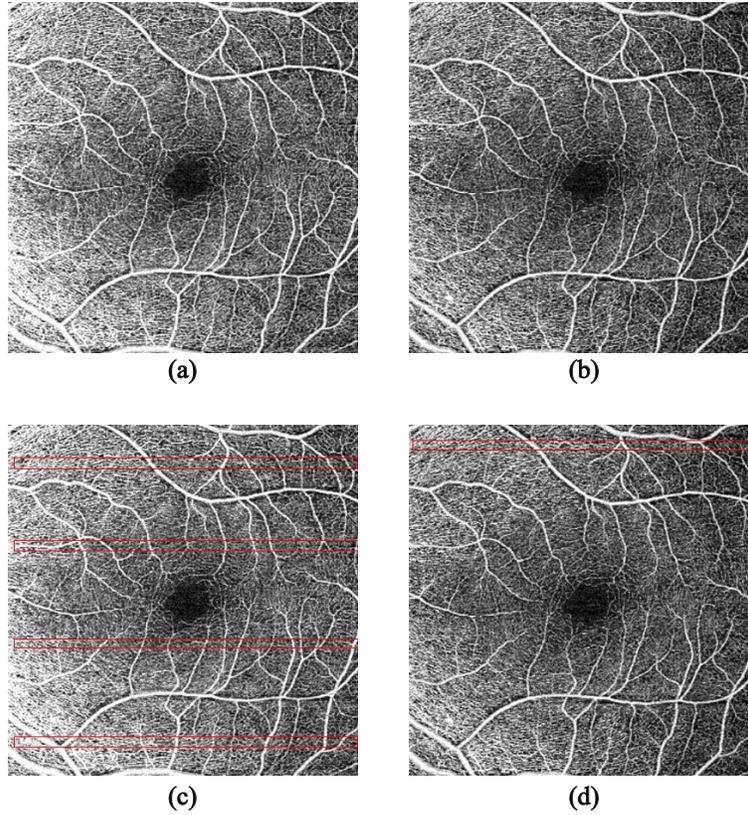


Figure 2.1: (a) and (b) are two images taken of the same eye at different timestamps. (c) and (d) consist of the annotation of motion artifacts in (a) and (b) respectively by the optometrist.

2.3.2 Image Registration

As already mentioned, the two images taken at different timesteps consists of disparity due to the natural motion of the eye. In order to align the images, we first extract the features from the two annotated images individually. Then, we match the features extracted from these images. We just take the top 20% of the matched features into consideration. Finally, we obtain an aligned image using homography[12] from the matched features. Fig 2.2 shows an example of aligned image.

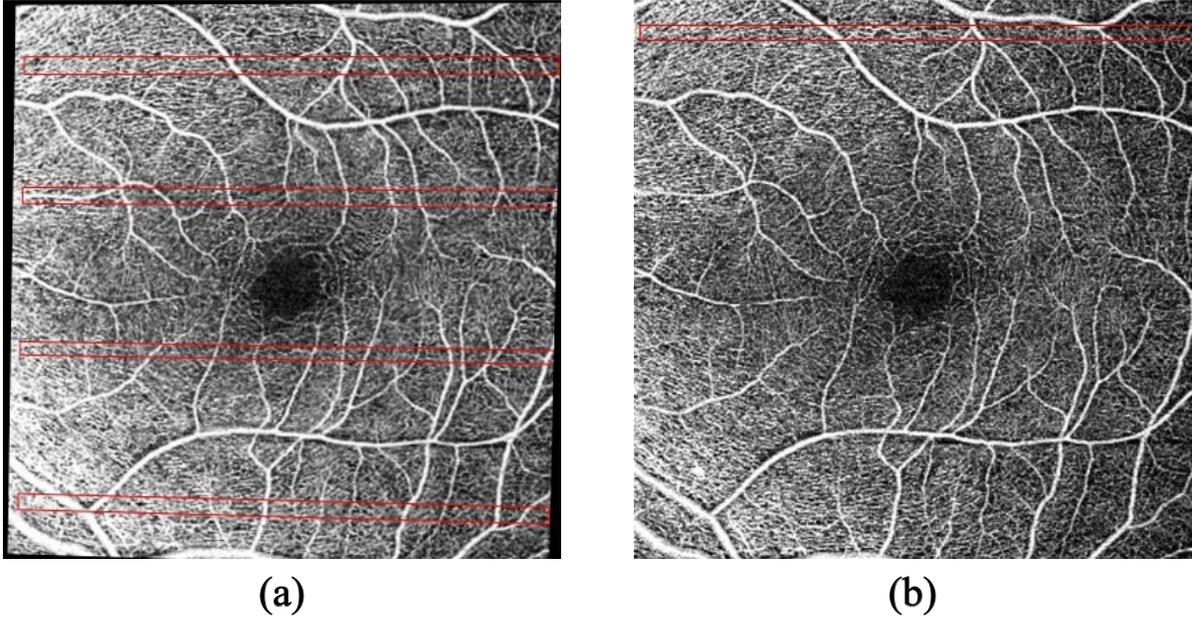


Figure 2.2: (a) is now aligned with (b) using homography[12]

2.3.3 Generating faux motion artifacts

Upon aligning the images using homography[12], we have two annotated images that are aligned to each other. The second image is the original image with no changes and the first image is aligned to the first image. Then, both the images are cropped from the same region along their width. This technique helps to produce more number of images for training as well as we discard the cropped image pairs if it consists on real motion artifact that were annotated by the optometrists.

After obtaining the two cropped images with no motion artifacts in them, we randomly select a thin strip from the second image. The height of this thin strip varies from 3 to 8 pixels. This thin strip is cropped and pasted onto the first image exactly at the same location it was extracted from in the second image. Thus, we obtain finally obtain a pair of images, with and without the motion artifact, which would be served as input to the model. While performing this process, we do not restrict ourselves to just one artifact per pair. We randomly introduce two artifacts in a single pair as well for the model to efficiently learn removing the artifacts if multiple of them are close enough. This step is pictorially represented in Fig 2.3.

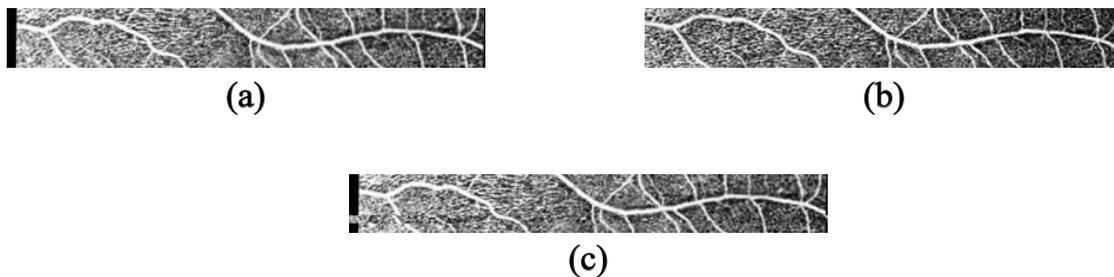


Figure 2.3: (a) and (b) are two images cropped from the Fig 2.2. (c) shows how a thin strip is taken from (b) and pasted in (a), replicating a motion artifact. (c) consisting of motion artifact is served as one of the inputs to the model and (a) acts as a ground truth image with no artifact.

2.3.4 Cardinal mask

The generation of cardinal mask plays the most vital role in the removal of motion artifacts. Initially, we start by taking a mask along the strip in the synthetically generated image. Then, we perform Otsu's Thresholding [41] on the horizontal image crops in order to retrieve the most relevant parts of the image and suppress the motion artifacts. Finally, we multiply the Otsu's generated image with the cardinal mask which allows to obtain the mask along the most relevant part of the motion artifact region. Fig 2.4 shows how the cardinal mask looks like and generated.



Figure 2.4: (a) shows the result after using Otsu's Thresholding [41]. (b) is the mask highlighting the region where the motion artifact exists. (c) is produced by multiplying (a) with (b).

2.3.5 Working

In this section, we delineate the inputs provided to the model and the anticipated outcome. Two inputs are provided to the network. ζ is defined as the image consisting of the artificially generated motion artifact, also shown in (c) of Fig 2.3. ϱ denotes the cardinal mask highlighting the relevant portions to the motion artifact to be focused on, also shown in (c) of Fig 2.4. Γ represents the network used.

$$\hat{\psi} = \Gamma([\zeta, \varrho]) \quad (2.1)$$

2.3.6 Loss Function

The original image ψ which was tampered to be introduced with artificial motion artifact serves as the ground truth image, shown in (a) of Fig 2.3. The L_{img} is mathematically represented as:

$$L_{img} = |\psi - \hat{\psi}| \quad (2.2)$$

Subsequently, in order for the model to specifically focus on motion artifact region, we apply the mask loss.

$$L_{mask} = |\psi * \varrho - \hat{\psi} * \varrho| \quad (2.3)$$

2.4 Experiments

2.4.1 Dataset

We assess the performance of the proposed training method in 140 images taken from 70 eyes. For every 70 eyes two images were taken after different timestamps. Each image is of the resolution 410×410 . These images were collected in Sankara Nethralaya, Chennai, India. Foveal microvascular dimensions (Angioplex) were imaged using optical coherence tomography angiography (OCTA; Cirrus 5000 Angioplex; Carl Zeiss Meditec Inc., Dublin).

2.4.2 Setup

All the training was done conducted on a single Nvidia Titan V [Graphics processing unit \(GPU\)](#) with 12 gigabytes of memory. All reported results and training times are specific to this [GPU](#) configuration. We utilized the PyTorch library [44] for the implementation.

2.4.3 Quantitative Results

The proposed training method is used on UNet[48], ENet[43] and UNet++[70]. Here we show the quantitative results for architectures only along the stripe in the image consisting of motion artifacts.

Table 2.1: PSNR and SSIM results across various architectures on the proposed training method.

Architecture	PSNR \uparrow	SSIM \uparrow
ENet[43]	14.79	0.36
UNet++[70]	15.41	0.45
UNet[48]	16.68	0.51

Table 2.2: Training time and different architectures.

Architecture	Training Time (hours)
ENet[43]	9
UNet++[70]	6
UNet[48]	2.5

In table 2.1, we show the PSNR and SSIM of various architectures on the proposed training method. In table 2.2, we show the training time taken by the networks.

2.4.4 Qualitative Results

In this section, we show the the qualitative results of the proposed method. In Fig 2.5, we can see how different architectures overcome the motion artifacts in the image that are supplied as input to them. The red arrow highlights those regions in the image that consist of the motion artifact. Green arrow highlights the resolved part, removing the artifact.

2.5 Conclusion

OCTA images are used for the diagnosis of retinal diseases. However, different kinds of artifacts are commonly found in them. In this work, we propose a training method which

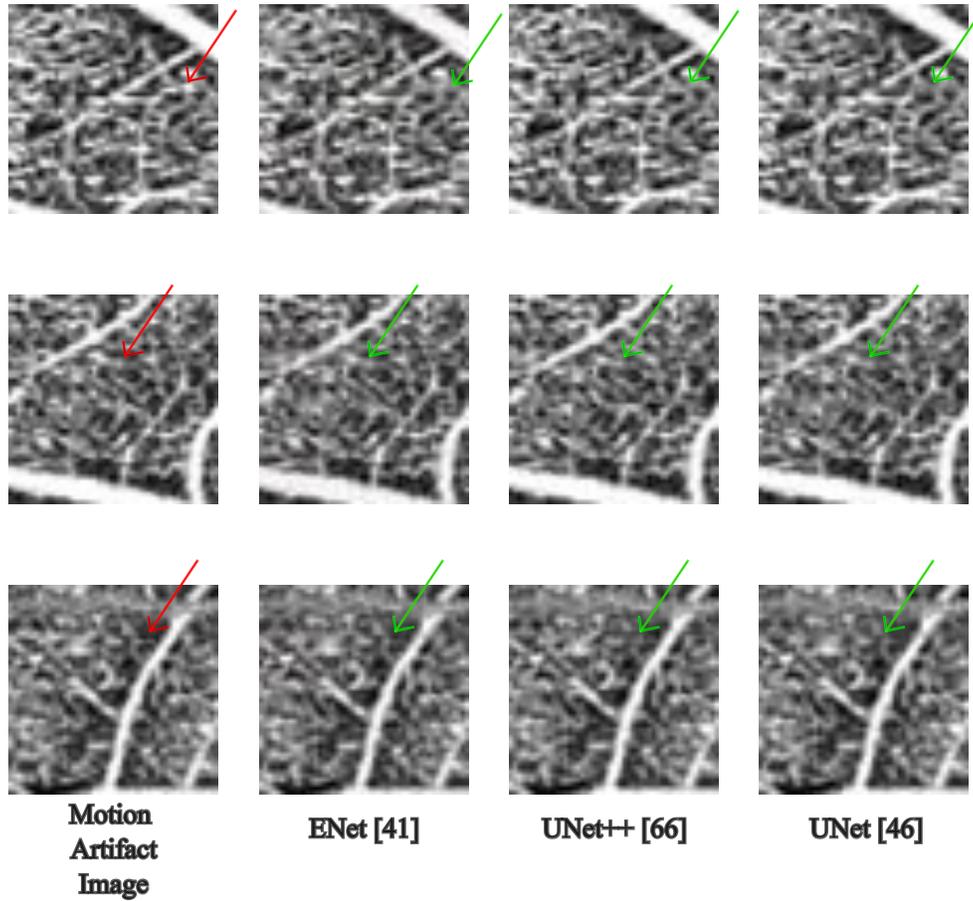


Figure 2.5: Red arrow shows the part of the image consisting of motion artifact. Green arrow shows how the architecture removes the motion artifact.

removes motion artifacts from the [OCTA](#) images. First, the annotation of the motion artifacts is performed by the optometrists. Then, the annotated images are aligned using homography[12]. After the alignment, we generate the artificial motion artifacts in the image which remarkably replicate the real ones. Finally, we create a cardinal mask which plays a crucial role in the removal of motion artifacts in the image. Since this training method can be used with multiple architectures, it becomes a plug-and-play system to overcome these artifacts.

Chapter 3

Super-Resolution of the OCTA image

This chapter is based on:

1. R Bhardwaj, J J Balaji, and V Lakshminarayanan. "OW-SLR: Overlapping Windows on Semi-Local Region for Image Super-Resolution". In MDPI: Journal of Imaging 2023, Vol. 9(11), p.246

3.1 Overview

There has been considerable progress in implicit neural representation to upscale an image to any arbitrary resolution. However, existing methods are based on defining a function to predict the RGB value from just four specific loci. Relying on just four loci is insufficient as it leads to losing fine details from the neighboring region(s). We show that by taking into account the semi-local region leads to an improvement in performance. In this paper, we propose applying a new technique called **Overlapping Windows on Semi-Local Region (OW-SLR)** to an image to obtain any arbitrary resolution by taking the coordinates of the semi-local region around a point in the latent space. This extracted detail is used to predict the RGB value of a point. We illustrate the technique by applying the algorithm to the OCTA images and show that it can upscale them to random resolution. This technique outperforms the existing state-of-the-art methods when applied to the OCT500 dataset. OW-SLR provides better results for classifying healthy and diseased retinal images such as diabetic retinopathy and normals from the given set of OCTA images. The project page is available at <https://rishavbb.github.io/ow-slr/index.html>.

3.2 Introduction

The primary objective of **super-resolution (SR)** is to obtain a credible **HR** image from a **LR** image. The major challenge is to retrieve the information which is too minute or almost non-existent, and to extrapolate this information to higher dimensions which is plausible to the human eye. Furthermore, the availability of paired **HR-LR** image data poses another concern. Typically, an image is downsampled using a specific method in the hope of encountering a real-life **LR** image that is somewhat similar. The aim of **SR** models is to fill in the deficient information between the **HR** and **LR** images, thereby bridging the gap. Also, for high-dimensional inputs like videos and 3D scans there are quite a few works in the literature [55, 58, 34, 59, 33, 16].

Most of the architectures [36, 53, 67, 45, 37] proposed for **SR** of images upsample them by a fixed factor only. This means that a separate architecture needs to be trained for each unseen upscaling factor. However, the real world is continuous in nature, whereas images are represented and stored as discrete values in 2D arrays. Inspired by [50, 42, 40, 54] for 3D shape reconstruction using implicit neural representation, [8] proposed **Local Implicit Image Function (LIIF)** to represent images in a continuous fashion. Some postprocessing is performed to obtain the **RGB** value of the query point. This approach enables representing

and manipulating images in a continuous manner, departing from the traditional discrete representation in 2D arrays.

In our work, we draw partial inspiration from advancements in 3D shape reconstruction, but we extend the approach by considering a semi-local region rather than relying solely on four specific locations. Our method allows for extrapolation to any random upscaling factor using the same architecture. This architecture takes into account the semi-local region and specifically learns to extract important details related to a query point in the latent space that needs to be upscaled. In this paper, we propose an image representation technique called Overlapping Windows for Semi-Local Representation in a continuous domain and we fine our work as follows: (i) Each image is represented as a set of latent codes, establishing a continuous nature. To determine the RGB value of a point in the HR image within the latent space, we employ a decoding function. (ii) This semi-local region is fed into network as input which generates the embeddings of the intricate details in it which have high probability of getting lost when an entire image is taken into consideration by the networks. (iii) The overlapping window technique allows for effective learning of features within the semi-local region around a point in the latent space using the embeddings. (iv) A decoder takes the features derived from the overlapping window technique and produces the RGB value of the corresponding point in the HR image.

In summary, our work makes two key contributions. Firstly, we introduce a novel technique called overlapping windows, which enables efficient learning of features within the semi-local region around a point. This approach allows for more effective representation and extraction of important details. Secondly, our architecture is capable of upscaling an image to any arbitrary factor, providing flexibility and versatility without the need for separate architectures for different upscaling factors. This contribution enables seamless and consistent image upscaling using a unified framework.

3.3 Related Work

During the early stages of SR research, images were typically upsampled by a certain factor using simple interpolation techniques, and the network was trained to learn the extrapolation of the LR images [10, 22]. However, this approach presents some issues. Firstly, the pre-upsampling process introduces more parameters compared to the post-upsampling process. Pre-upsampling is defined as upscaling the input image and then passing it through the network, whereas post-upsampling is defined as passing the image through the network and then upscaling the feature map. Secondly, due to the higher requirement of parameters more training time becomes a requisite. The network needed

to learn the intricacies of the pre-upsampling method, which added to the overall training complexity. Finally, the pre-upsampling process using traditional bicubic interpolation does not yield realistic results during testing. Since it is the first step of the **SR** pipeline, the network often attempts to mimic this interpolation, which limits the realism of the output images. On the other hand, post-upsampling approaches, where the **LR** image is downscaled in the very first step, typically involve the use of bicubic interpolation for resizing. However, downscaling an image, even with bicubic interpolation, tends to yield more realistic results compared to upscaling. As a result, the research focus has shifted towards post-upsampling techniques, which provides more efficient and realistic **SR** results by leveraging downscaling with appropriate interpolation methods in the very first step.

As already mentioned, downscaling of images happens as the initial step in post-upsampling process. The network learns features from the downscaled image and the upsamples the learned features towards the very end. A technique proposed by Shi et al. [53] in their work is known as sub-pixel convolution. Sub-pixel convolution handles the extrapolation of each pixel by accumulating the features along the channel of that pixel. By rearranging the feature channels, sub-pixel convolution enables the network to effectively upscale the **LR** image to a higher resolution. While sub-pixel convolution provides a practical solution for upsampling by integral factors ($\times 1$, $\times 2$, $\times 3$, etc.), it does not support fractional upsampling factors ($\times 1.4$, $\times 2.9$, etc.). However, for cases where fixed integral upsampling factors are sufficient, sub-pixel convolution offers an efficient approach to achieving high-quality upsampling. The work by Ledig et al. [29] introduced the use of multiple residual blocks for feature extraction in **SR** tasks. Their approach demonstrated the effectiveness of residual blocks in capturing and enhancing image details. Building upon Ledig et al.’s work, Lim et al. [37] proposed an enhanced **SR** model that incorporated insights regarding batch normalization. They postulated that removing batch normalization from the residual blocks could lead to improved performance for **SR** tasks. This is because batch normalization tends to normalize the input, which may reduce the network’s ability to capture and amplify the fine details required for **SR**. Removing batch normalization not only results in a reduction in memory requirements but also makes the network faster. Additionally, the work by Shi et al. [53] contributed to the development of various approaches for **SR** using **Convolutional Neural Networks (CNNs)**. These approaches include methods proposed by [67, 29, 39, 68]. These methods aimed to enhance feature extraction capabilities specifically tailored for **SR** problems, further advancing the state-of-the-art in **SR** research.

After the success of **CNNs** in **SR** tasks, researchers explored the use of **Generative Adversarial Networks (GANs)** to further improve **SR** performance. Several works, such as [29, 51, 63], introduced different **GAN** architectures for extrapolating **LR** images to

higher resolution. [Enhanced Super-Resolution Generative Adversarial Network \(ESRGAN\)](#) proposed by Wang et al. [64] introduced a perceptual loss function and modified the generator network to produce [HR](#) images. This perceptual loss function aimed to align the visual quality of the generated [HR](#) images with that of the ground truth [HR](#) images, improving the perceptual realism of the results.

In Real-ESRGAN [62], the authors addressed the issue of using [LR](#) images downsampled with simple techniques like bicubic interpolation during training. They note that real-world [LR](#) images undergo various types of degradations, compressions, and noise, unlike the simple interpolation-based downsampling. To simulate realistic [LR](#) images during training, they proposed a novel technique that subjected the training images to various degradation processes, mimicking real-life scenarios. Additionally, Real-ESRGAN[62] introduced an U-Net[48] discriminator to enhance the adversarial training process and improve the quality of the generated [HR](#) images.

3.4 Method

We illustrate the three main components of our approach in this section along with its pictorial representation in Figure 3.1. In Section 3.4.1, we introduce the backbone of our framework. We represent the [LR](#) image as a feature map, which serves as the basis for subsequent processing and analysis. In Section 3.4.2, we demonstrate how we find the semi-local region of an arbitrary point in the [HR](#) image. This region contains valuable information that helps determine the corresponding [RGB](#) value. In Section 3.4.3, we highlight the Overlapping Windows technique, which plays a crucial role in predicting the [RGB](#) value of a point in the [HR](#) image. We accomplish this by leveraging the semi-local region extracted around the sampling points of the feature map. These three parts collectively form the foundation of our approach, allowing for accurate prediction of [RGB](#) values.

3.4.1 Backbone Framework

To extract features from the [LR](#) image, we employ the [Enhanced Deep Residual Networks \(EDSR\)](#) [37]. Specifically, we utilize the baseline architecture of [EDSR](#)[37], which consists of 16 residual blocks.

$$\psi = EDSR(I_{LR}) \tag{3.1}$$

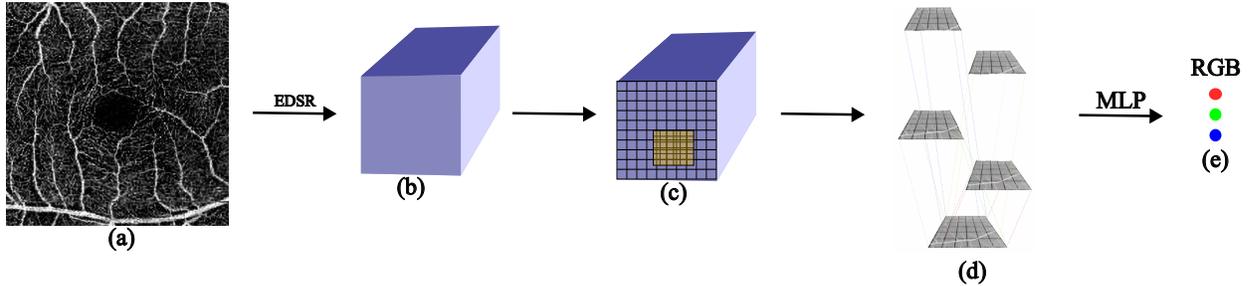


Figure 3.1: (a) An LR image is taken. (b) It is passed through EDSR [37] and a feature map is produced. (c) Locating the semi-local region ($M = 6$) around a random selected point from HR image. (d) Semi-local region is passed through the proposed Overlapping Windows. (e) This output is passed through the MLP to give out the RGB value of a randomly selected point. Steps (c-e) are performed for all the points in the HR image.

Given an LR image denoted as $I_{LR} \in \mathbb{R}^{H \times W \times C}$, we express it in the form of a feature map $\psi \in \mathbb{R}^{P \times Q \times D}$. Here, H and W represent the height and width of the LR image, respectively, and C signifies the number of channels. P and Q represent the spatial dimensions of the feature map, and D denotes the depth of the feature map.

3.4.2 Locating the Semi-Local Region

In our scenario, we aim to predict the RGB value at any random point in a continuous HR image of arbitrary dimensions. Let $I_{HR} \in \mathbb{R}^{X \times Y \times C}$ represent the HR image. To predict the RGB value at a specific point, we first select a point of interest. Then, we identify its corresponding spatially equivalent point in the feature map ψ obtained from the LR image using bilinear interpolation denoted as \mathcal{U}_{BI} .

$$\hat{\mathbf{x}} = \mathcal{U}_{BI}(x, \psi) \quad (3.2)$$

where $\hat{\mathbf{x}}$ and x are the 2D coordinates of the ψ and I_{HR} respectively.

Furthermore, we extract a square semi-local region around this corresponding point. The size of this region is determined by a length parameter M units, where each unit dimension of the square region corresponds to the inverse of the dimensions P and Q of the feature map ψ along its length and breadth respectively defined in Equation (5) which is used to find the discrete positions in the semi-local region. Once we have identified the square semi-local region around the corresponding point in the feature map ψ , we proceed

to extract $M \times M$ depth features from this region using Equation (3). These depth features capture the important information necessary for predicting the RGB value at the desired point in the HR image. To extract these features, we employ a closest Euclidean distance approach denoted by \mathfrak{d}_{ED} . Each point within the $M \times M$ region in ψ is mapped to the nearest point in the latent space, which represents the extracted depth feature. Figure 3.2 illustrates the working of selecting of features from the feature map. This mapping ensures that we capture the most relevant information from the semi-local region.

$$\hat{\mathbf{X}} = (\hat{\mathbf{x}}_x - \psi_x * i, \hat{\mathbf{x}}_y - \psi_y * j) \quad (3.3)$$

$$i = \left\{ \frac{-M}{2}, \frac{-M}{2} + 1, \dots, \frac{+M}{2} - 1, \frac{+M}{2} \right\}, j = \left\{ \frac{-M}{2}, \frac{-M}{2} + 1, \dots, \frac{+M}{2} - 1, \frac{+M}{2} \right\} \quad (3.4)$$

$$\psi_x = \frac{1}{P}, \psi_y = \frac{1}{Q} \quad (3.5)$$

Thus, $\hat{\mathbf{X}}$ holds the 2D coordinates of all the $M \times M$ points.

$$S = \mathfrak{d}_{ED}(\hat{\mathbf{X}}, \psi) \quad (3.6)$$

Figure 3.3 illustrates how the semi-local region is identified and used to extract the $M \times M$ depth features from the feature map ψ . This depiction helps to visualize the steps involved in the feature extraction process.

3.4.3 Overlapping Windows

After extracting the semi-local region $S \in \mathbb{R}^{M \times M \times D}$, our objective is to obtain the RGB value of the center point using this region. To achieve this, we employ a overlapping window-based approach. We start with four windows, each with a size of $M - 1$, positioned at the four corners of S . Each window extracts information from its respective region and passes it on to the next subsequent window in the process. With each iteration, the size of the window decreases by 1 until it reaches a final size of $\frac{M}{2}$. This iterative process ensures that information is progressively gathered and refined towards the center point. This approach allows us to effectively capture and utilize the information from the semi-local region while focusing on the features that are most relevant for determining the RGB value.

$$\Gamma = s_i * w_i \quad (3.7)$$

In each iteration i , where the window size decreases by 1 for the next step, we utilize weights w_i for combining the features from all four corners. This ensures that the information from each corner is properly incorporated and made available for the subsequent

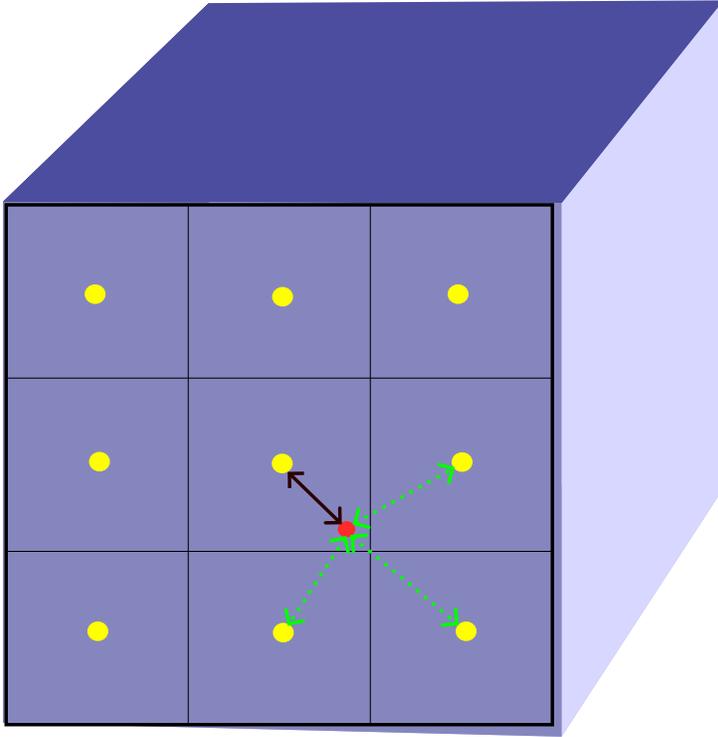


Figure 3.2: To extract features from a feature map of size 3×3 , we focus on a specific query point represented by a red dot. In order to determine which pixel locations in the feature map correspond to this query point, we compute the Euclidean distance between the query point and the center points of each pixel location. In the provided image, the black line represents the closest pixel location in the feature map to the query point.

iteration. In the last step, we take a final window size of 2, but instead of being positioned at the corners as in previous iterations, it is centered around the target point of interest. The features extracted from this final window are then passed through a [MLP](#) to make the final prediction.

By adapting the window positions and sizes throughout the iterations, we effectively capture and aggregate the relevant information from the semi-local region. This approach allows us to make accurate predictions at the target point, utilizing the combined features from all iterations and the final [MLP](#)-based processing. Figure 3.4 shows the working of the overlapping windows.

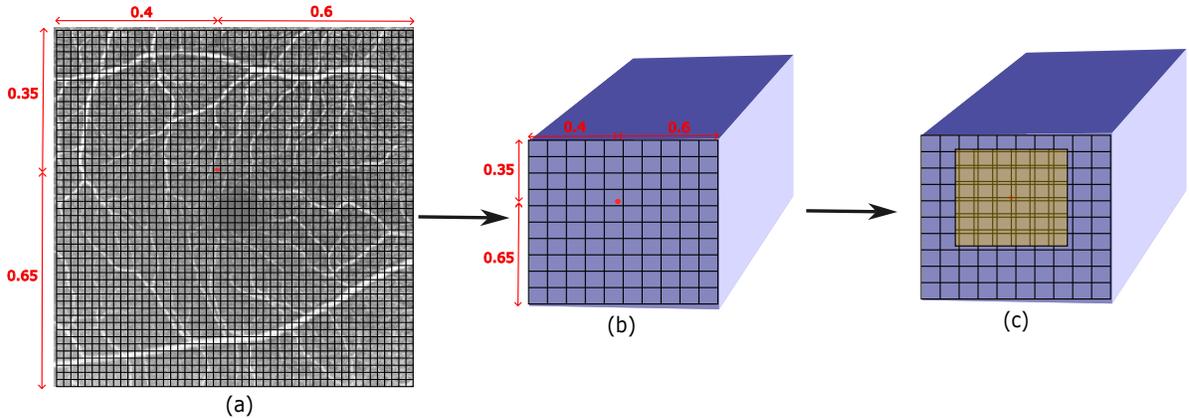


Figure 3.3: (a) Given an HR image, a point of interest (red dot) is selected to predict its RGB value. (b) Its corresponding spatially equivalent 2D coordinate is selected from the feature map. (c) Locating the semi-local region ($M = 6$) around the calculated 2D coordinate.

3.5 Results and Discussion

3.5.1 Dataset

We used the OCT500 [35] dataset and randomly sampled 524 images from it to train our network. It consists of 300 3×3 OCTA images and 224 6×6 OCTA images. For evaluation, 80 images were selected and we report the results using PSNR metric.

3.5.2 Implementation Details

During the training process, we apply downsampling to each image using bicubic interpolation in PyTorch [1]. This downsampling is performed by selecting a random factor, which introduces the desired level of degradation to the images. For training, we utilize a batch size of 16 images. From each HR image, we randomly select 1500 points for which we aim to calculate the RGB values. These points serve as the targets for our network during the optimization process.

To optimize the network, we employ the L1 loss function and use the Adam optimizer [23]. The learning rate is initialized as $1.e - 4$ and is decayed by a factor of 0.3 at specific epochs, namely [40, 60, 70]. We train the network for a total of 100 epochs, allowing it to learn the necessary representations and refine its predictions over time.

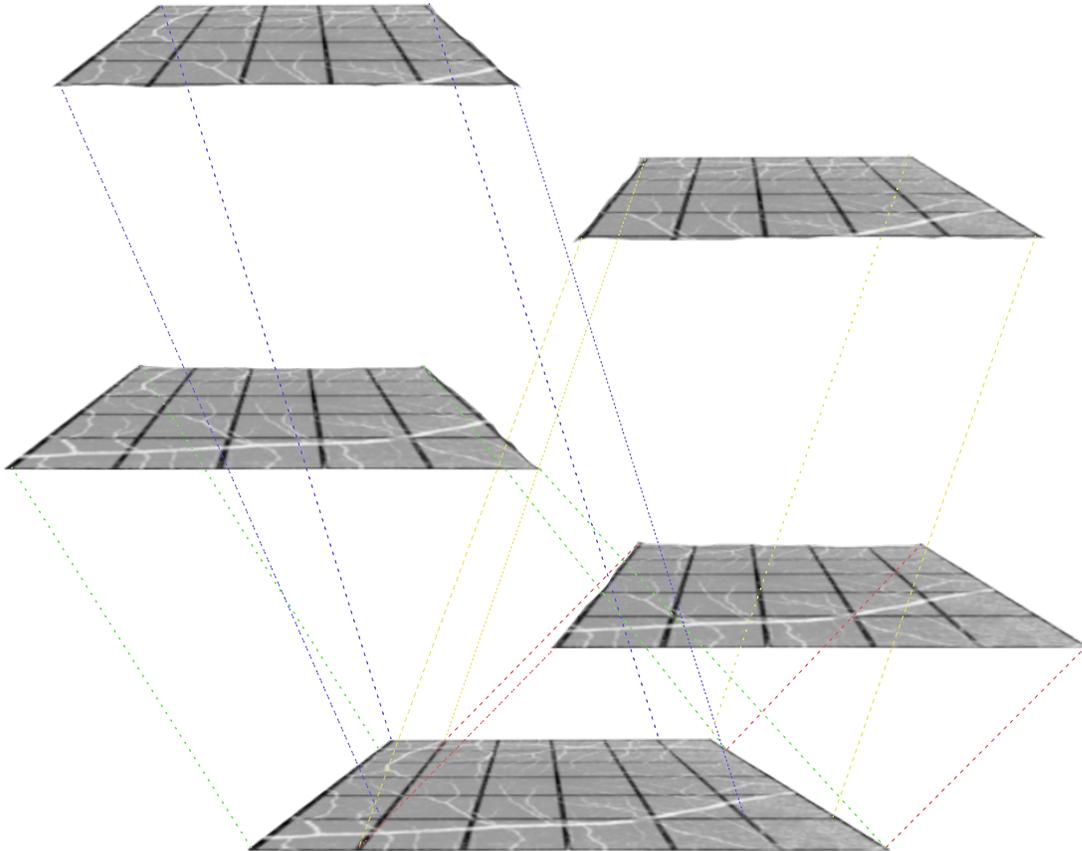


Figure 3.4: The first iteration of overlapping windows, where the window size = $M - 1$ ($M = 6$). Assuming the feature map is of negligible depth and four windows are positioned at the four corners of the feature map.

Furthermore, each **LR** image is converted into a feature map of size 48×48 with a depth of 64 using the **EDSR**-baseline architecture. This conversion process ensures that the **LR** images are properly represented and aligned with the architecture used in the training process.

3.5.3 Quantitative Results

In Figure 3.5, we present a comparison of the performance of our proposed **OW-SLR** method against existing works. The original image patch is first downsampled using bicubic

interpolation to a lower resolution. It is evident that there is a significant loss of image quality in the **LR** patches compared to the **Ground Truth (GT)** image. However, our model outperforms the other existing methods, demonstrating a significant improvement when the **LR** image is extrapolated to a higher scale. The results obtained by our model show better preservation of details and higher fidelity compared to the other approaches when the given image is extrapolated to higher scale. The **PSNR** results of each image are shown in Table 3.1.

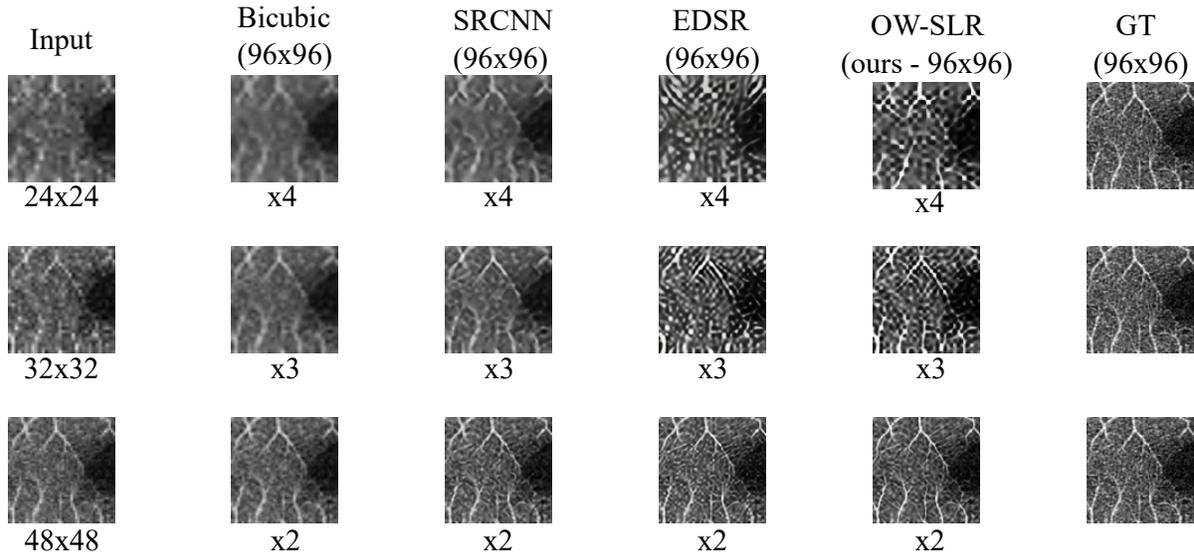


Figure 3.5: A 96×96 patch is taken and its size is reduced to 24×24 (first row), 32×32 (second row) and 48×48 (third row) using bicubic interpolation. Our architecture uses the same set to weights reproduce the given results. However, others require different set of weights for a newer scale to be trained on. The **PSNR** results of each image are shown in Table 3.1.

It is worth noting that our model achieves these results for different scaling factors using the same set of weights trained once. In contrast, the other models would need to be retrained for each new scale to which the **LR** image is extrapolated. This highlights the versatility and efficiency of our model in handling various scaling factors without the need for additional training.

In Table 3.2, we provide the upscaling time taken by the proposed model by different factors, while training it just once.

In Table 3.3, we present the results of this technique compared to the existing state-of-the-art methods on the OCT500 [35] dataset. The evaluation metric used in this case is the

Table 3.1: PSNR result of each of the input images across different methods shown in Figure 3.5.

Patch Size	Bicubic	SRCNN [10]	EDSR [37]	OW-SLR (Ours)
24×24	11.96	12.87	13.79	13.92
32×32	14.18	15.10	16.04	16.26
48×48	15.37	16.89	17.66	17.98

Table 3.2: Time taken to extrapolate a 320×320 image on a single Nvidia Titan V of 12 Gigabyte size.

Extrapolation Factor	Time Taken (In Seconds)
$2\times$	6.48
$2.4\times$	8.90
$3\times$	12.01
$3.9\times$	19.46
$4.5\times$	26.29
$5\times$	33.75

PSNR. Our work demonstrates superior performance compared to LIIF, highlighting the effectiveness of considering the semi-local region instead of solely focusing on four specific locations. By incorporating the information from the semi-local region, our approach achieves improved results in terms of PSNR, showcasing the benefits of our methodology for super-resolution tasks.

3.6 Conclusion

OCTA images help us for the diagnosis of retinal diseases. However, due to various reasons like speckle noise, movement of the eye, hardware incapacibilities, etc. we lose onto intricate details in the capillaries that play a crucial role for correct diagnosis. We propose this architecture which upscales a given LR image to arbitrary higher dimensions with enhanced image quality. First, we extract the image features using a backbone architecture. We then select a random point in the HR image and calculate its equivalent spatial point in the extracted feature map. We find the semi-local region around this calculated point and pass

Table 3.3: PSNR result on the 300 images from OCT500 [35]

Methods	PSNR \uparrow
Real-ESRGAN [64]	15.66
SRCNN [10]	16.51
EDSR [37]	17.49
LIIF [8]	17.60
OW-SLR (ours)	17.93

it through the proposed Overlapping Windows architecture. Finally, an MLP is used to predict the RGB value using the output of the overlapping window architecture. We hope our work will help the people in the medical field in their diagnosis. PSNR 17.93 is achieved for the OCT500 dataset which outperforms the other state-of-the-art work. The technique outperforms the existing methods and allows upscaling images to arbitrary resolution by training the architecture just once.

While effective, it is worth noting that this algorithm does come with a slightly higher computational cost due to its consideration of the semi-local region. There remains potential for further enhancements in both computational efficiency and accuracy while taking the semi-local region into account. This work will provide a stepping stone for future researchers to make strides in this direction.

Chapter 4

Evaluation of Clinical Parameters

4.1 Introduction

OCTA images are used for the diagnosis of retinal diseases. In Chapter 2, we remove one of the most difficult artifacts - motion artifacts - from the OCTA image. In Chapter 3, we take the image, free from motion artifacts and increase the dimensions of the image to allow better diagnosis by the medical practitioners using super-resolution. In this chapter, we will be evaluating if this higher dimension image is distorting the clinical parameters or maintaining the similar trend.

4.2 Software

To perform the evaluation process, we use the software FAZSeg[60] which gives 15 parameters related to the FAZ in the OCTA image. The 15 parameters that FAZSeg[60] outputs are (i) Area, (ii) Diameter, (iii) Major Axis Length, (iv) Minor Axis Length, (v) Perimeter, (vi) F_{min} , (vii) F_{max} , (viii) Inner Circle Radius, (ix) Eccentricity, (x) Circum Circle Radius, (xi) Orientation, (xii) Tortuosity, (xiii) Vessel Diameter Index, (xiv) Circularity and (xv) Vessel Diameter Index. The user interface of the FAZSeg[60] software can be seen in Fig 4.1.

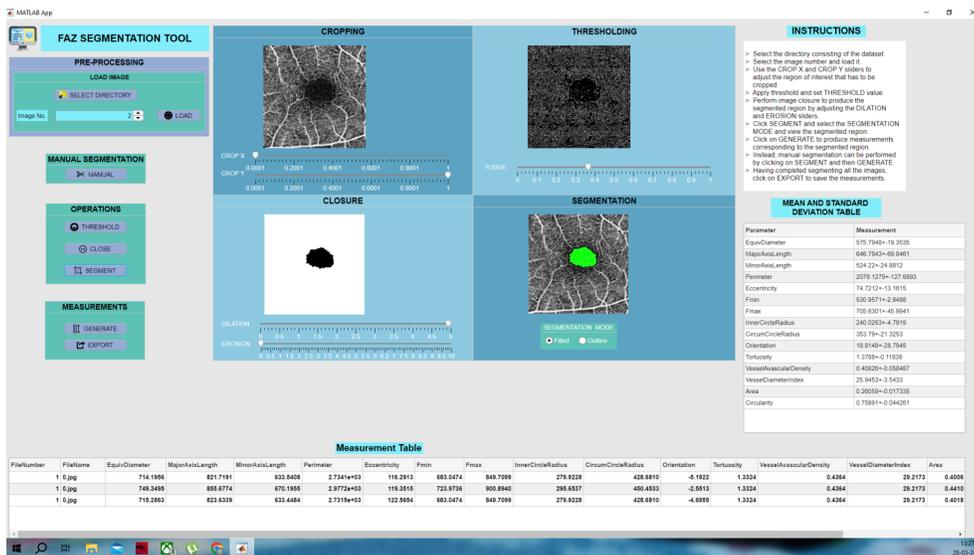


Figure 4.1: FAZSeg[60] software being used to find the 15 parameters.

4.3 Data

For the entire evaluation process we use the FAZID[2] dataset. This dataset consists of diabetic, normal and myopia images. For our evaluation process we use 10 normal images.

4.4 Experiment

As already mentioned, we take 10 normal images for evaluation. We remove the motion artifacts from the images and upscale each image by $\times 2$ and $\times 3$, thus now having 30 images in total. The FAZSeg[60] software is able to accept only 420×420 images. We manually crop the upscaled images to size 420×420 centered in the FAZ, since that is the region of interest to get the 15 parameters. In order to maintain the consistency of results rendered by the FAZSeg[60] software, we run each image four times through the software and then take an average of all the parameters.

After averaging the parameters for each image, we then move on to average the parameters for all the 10 images to understand the trend of the raw image and the extrapolated image. The table 4.1 shows the averaged values across the 10 images. A p-value < 0.05 was considered significant for all statistical tests.

4.5 Result and Discussion

All statistical analyses were performed using SPSS version 20[57]. Table 4.1 shows the final results of impact of FAZ dimensions before and after applied super resolution. Out the 15 different FAZ parameters, except the circularity index, orientation, tortuosity all other dimensions showed significant difference ($p = < 0.05$) after the super resolution.

In order to check the linear relationship between FAZ dimension before and after the super resolution, a correlation plot was done for the common and widely used parameters FAZ's area, perimeter and circularity index. The results of correlation showed in figures 4.2, 4.3 and 4.4. All three parameters showed a positive and linear with all three parameter. However, the area and perimeter had an asymmetrical correlation co-efficient values between $\times 2$ and $\times 3$. The highest correlation co-efficient ($r^2 = > 0.9$) was seen between $\times 1$ and $\times 2$ in circularity index and the least correlation co-efficient ($r^2 = < 0.21$) was seen between $\times 1$ and $\times 3$ in area. These variation possibly due FAZSeg[60] application limitation.

Table 4.1: A comparison of FAZ parameters before and after super-resolution.

	Raw ($\times 1$)	Image ($\times 2$)	($\times 3$)	p -Value ^{&}
Area (mm^2)	0.32 ± 0.07	1.08 ± 0.23	2.53 ± 0.45	< 0.001
Circularity Index \blacktriangle	0.75 ± 0.06	0.75 ± 0.07	0.76 ± 0.06	0.166
Circumcircle radius (mm)	382.59 ± 47.17	714.16 ± 94.77	1083.22 ± 89.17	< 0.001
Diameter (mm)	638.39 ± 70.92	1168.13 ± 126.58	1788.33 ± 161.72	< 0.001
Eccentricity (mm)	84.72 ± 24.68	191.53 ± 40.69	302.45 ± 49.64	< 0.001
F_{max} (mm)	761.91 ± 95.46	1426.24 ± 189.95	2158.94 ± 176.64	< 0.001
F_{min} (mm)	605.83 ± 65.96	1104.35 ± 134.8	1672.56 ± 175.95	< 0.001
Inner circle radius (mm)	264.61 ± 25.49	477.68 ± 56.11	744.61 ± 97.53	< 0.001
Major axis length (mm)	718.04 ± 88.23	1336.15 ± 177.52	2021.5 ± 189.87	< 0.001
Minor Axis Length (mm)	580.2 ± 61.33	1051.29 ± 113.59	1627.76 ± 179.19	< 0.001
Orientation ($^\circ$)	-15.21 ± 14.09	-20.25 ± 15.67	-18.37 ± 12.67	0.301
Perimeter (mm)	2328.99 ± 340.34	4259.78 ± 598.54	6469.74 ± 653.35	< 0.001
Tortuosity \blacktriangle	1.35 ± 0.06	1.37 ± 0.08	1.4 ± 0.08	0.585
Vessel Avascular Density (VAD) \blacktriangle	0.42 ± 0.01	0.35 ± 0.01	0.28 ± 0.02	0.007
Vessel Diameter Index (VDI) (mm)	29.21 ± 0.73	39.55 ± 1.55	44.52 ± 0.77	< 0.001

mm :millimeters, \blacktriangle : dimensionless, $^\circ$: degree, &: Friedman Test

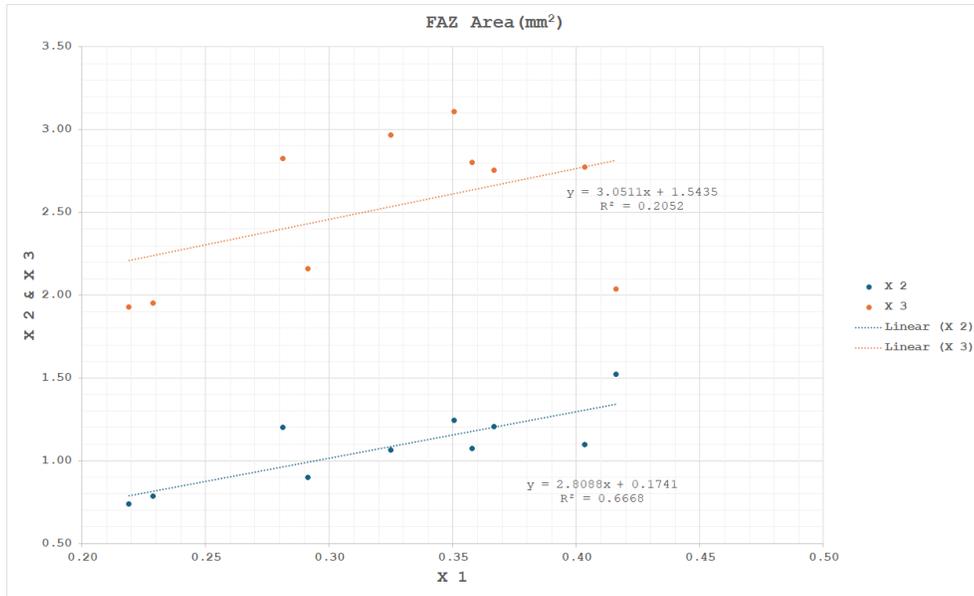


Figure 4.2: Correlation between FAZ area measured before and after super-resolution.

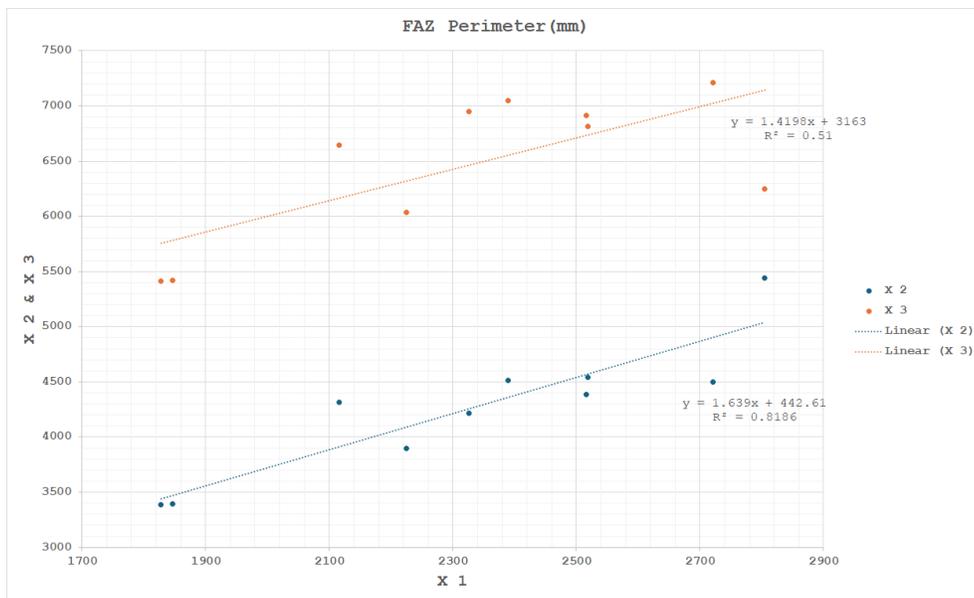


Figure 4.3: Correlation between FAZ perimeter measured before and after super-resolution.

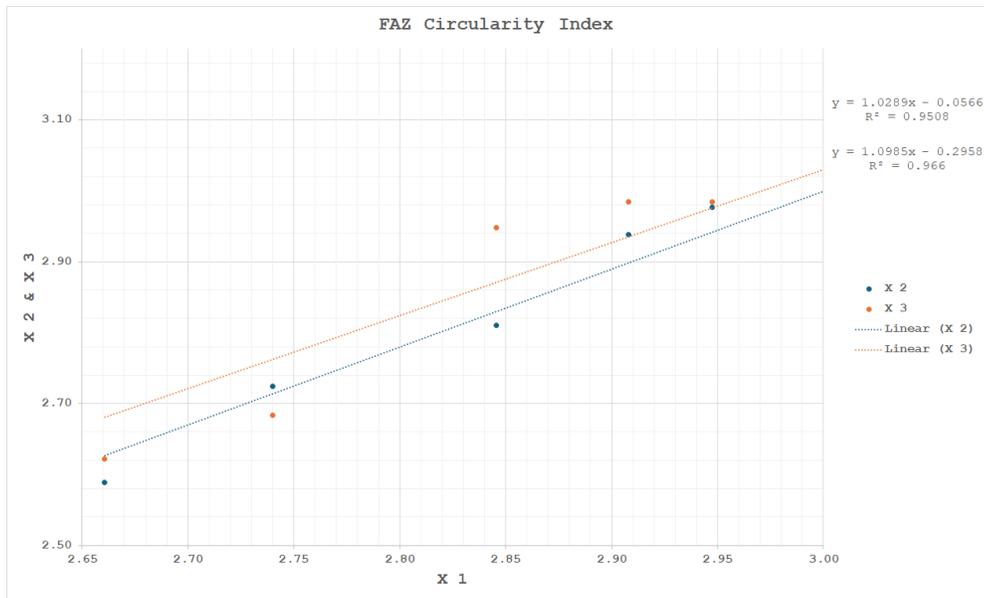


Figure 4.4: Correlation between FAZ circularity index measured before and after super-resolution.

Chapter 5

Conclusion and future research

OCTA images help us for the diagnosis of retinal diseases. In this work, we first start by removing one of the most difficult artifacts found in **OCTA** images - motion artifacts. Initially, optometrists annotate the motion artifacts present in the images. Subsequently, the annotated images undergo alignment using homography techniques. Following alignment, we simulate artificial motion artifacts in the images that closely resemble real ones. Lastly, we develop a cardinal mask essential for effectively removing motion artifacts from the images. This training methodology is adaptable across various architectures, making it a versatile solution for mitigating these artifacts.

After the obtaining an image free from motion artifacts, we perform super-resolution to enhance the quality of the image. To address this, we propose an architecture capable of upscaling **LR** images to higher dimensions with improved image quality. Initially, we extract image features using a backbone architecture. Subsequently, we identify a random point in the **HR** image and determine its corresponding spatial point in the extracted feature map. From there, we define a semi-local region around this point and process it through our proposed Overlapping Windows architecture. Finally, we employ a **MLP** to predict **RGB** values based on the output of the overlapping window architecture. We anticipate that our approach will benefit medical professionals in their diagnostic endeavors.

Regarding the drawbacks, it is a necessity to annotate the image every time to remove the motion artifacts. One of the ideas for the future research could be to develop a system which automatically annotates the motion artifacts. However, it would require a lot of images and skilled optometrists to generate the training images. In case of super-resolution, it is important to note that this algorithm carries a slightly higher computational overhead due to its consideration of semi-local regions. There is room for optimizing computational

efficiency without compromising accuracy while incorporating semi-local information. This research serves as a foundation for future advancements in this field, paving the way for more sophisticated algorithms and improved diagnostic capabilities in medical imaging.

References

- [1] P. Adam, G. Sam, M. Francisco, L. Adam, B. James, C. Gregory, K. Trevor, Z. Lin, G. Natalia, and A. Luca. Pytorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems*, 2019.
- [2] Arpit Agarwal, Rajiv Raman, Vasudevan Lakshminarayanan, et al. The foveal avascular zone image database (fazid). In *Applications of Digital Image Processing XLIII*, volume 11510, pages 507–512. SPIE, 2020.
- [3] Baumann B, Pircher M, Götzinger E, and Hitzenberger CK. Full range complex spectral domain optical coherence tomography without additional phase shifters. *Opt Express*.15(20):13375-87. PMID: 19550607; PMCID: PMC2978327., 2007.
- [4] B. Baumann. Polarization sensitive optical coherence tomography: A review of technology and applications. *Applied Sciences*, vol. 7, no. 5, p. 474, 2017.
- [5] H. G. Bezerra, M. A. Costa, G. Guagliumi, and et al. Intracoronary optical coherence tomography: A comprehensive review: Clinical and research applications. *JACC: Cardiovascular Interventions*, vol. 2, no. 11, pp. 1035-1046, 2009.
- [6] J. F. De Boer, C. K. Hitzenberger, and Y. Yasuno. Polarization sensitive optical coherence tomography-a review. *Biomedical Optics Express*, vol. 8, no. 3, pp. 1838-1873, 2017.
- [7] Bouma, Brett, Yun, Seok-Hyun, Vakoc, Benjamin, Suter, Melissa, Tearney, and Guillermo. Fourier-domain optical coherence tomography: Recent advances toward clinical utility. *Current opinion in biotechnology*, vol. 20, pp. 111-8, 2009.
- [8] Y. Chen, S. Liu, and X Wang. Learning continuous image representation with local implicit image function. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021.

- [9] WooJhon Choi, Eric M Moulton, Nadia K Waheed, Mehreen Adhi, ByungKun Lee, Chen D Lu, Talisa E de Carlo, Vijaysekhar Jayaraman, Philip J Rosenfeld, Jay S Duker, et al. Ultrahigh-speed, swept-source optical coherence tomography angiography in nonexudative age-related macular degeneration with geographic atrophy. *Ophthalmology*, 122(12):2532–2544, 2015.
- [10] C. Dong, C.C. Loy, K. He, and X Tang. Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2015.
- [11] Mariantonia Ferrara, Gaia Lugano, Maria Teresa Sandinha, Victoria R Kearns, Brendan Geraghty, and David HW Steel. Biomechanical properties of retina and choroid: a comprehensive review of techniques and translational relevance. *Eye*, 35(7):1818–1832, 2021.
- [12] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.
- [13] Alejandro F Frangi, Wiro J Niessen, Koen L Vincken, and Max A Viergever. Multiscale vessel enhancement filtering. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI’98: First International Conference Cambridge, MA, USA, October 11–13, 1998 Proceedings 1*, pages 130–137. Springer, 1998.
- [14] Simon S Gao, Yali Jia, Miao Zhang, Johnny P Su, Gangjun Liu, Thomas S Hwang, Steven T Bailey, and David Huang. Optical coherence tomography angiography. *Investigative ophthalmology & visual science*, 57(9):OCT27–OCT36, 2016.
- [15] Simon S Gao, Gangjun Liu, David Huang, and Yali Jia. Optimization of the split-spectrum amplitude-decorrelation angiography algorithm on a spectral optical coherence tomography system. *Optics letters*, 40(10):2305–2308, 2015.
- [16] R. Girshick, J. Donahue, T. Darrell, and J. Malik. Region-based convolutional networks for accurate object detection and segmentation. *IEEE Trans. Pattern. Anal. Mach. Intell.*, 2016.
- [17] Michel Goossens, Frank Mittelbach, and Alexander Samarin. *The L^AT_EX Companion*. Addison-Wesley, Reading, Massachusetts, 1994.
- [18] David Huang, Eric A Swanson, Charles P Lin, Joel S Schuman, William G Stinson, Warren Chang, Michael R Hee, Thomas Flotte, Kenton Gregory, Carmen A Puliafito, et al. Optical coherence tomography. *science*, 254(5035):1178–1181, 1991.

- [19] Yali Jia, Ou Tan, Jason Tokayer, Benjamin Potsaid, Yimin Wang, Jonathan J Liu, Martin F Kraus, Hrebesh Subhash, James G Fujimoto, Joachim Hornegger, et al. Split-spectrum amplitude-decorrelation angiography with optical coherence tomography. *Optics express*, 20(4):4710–4725, 2012.
- [20] Schuman JS. Spectral domain optical coherence tomography for glaucoma (an aos thesis). *Trans Am Ophthalmol Soc.* 2008;106:426-58. PMID: 19277249; PMCID: PMC2646438, 2008.
- [21] Amir H Kashani, Chieh-Li Chen, Jin K Gahm, Fang Zheng, Grace M Richter, Philip J Rosenfeld, Yonggang Shi, and Ruikang K Wang. Optical coherence tomography angiography: a comprehensive review of current methods and clinical applications. *Progress in retinal and eye research*, 60:66–100, 2017.
- [22] J. Kim, J.K. Lee, and K.M. Lee. Accurate image super-resolution using very deep convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.
- [23] D.P. Kingma and J.A. Ba. A method for stochastic optimization. *arXiv:1412.6980*, 2014.
- [24] S. Kishi. Impact of swept source optical coherence tomography on ophthalmology. *Taiwan Journal of Ophthalmology*, vol. 6, no. 2, pp. 58-68, 2016.
- [25] Donald Knuth. *The T_EXbook*. Addison-Wesley, Reading, Massachusetts, 1986.
- [26] Wang L, Fu R, Xu C, and Xu M. Methods and applications of full-field optical coherence tomography: a review. *J Biomed Opt.*27(5):050901. PMID: 35596250; PMCID: PMC9122094., 2022.
- [27] Leslie Lamport. *L^AT_EX — A Document Preparation System*. Addison-Wesley, Reading, Massachusetts, second edition, 1994.
- [28] Max WK Law and Albert CS Chung. Three dimensional curvilinear structure detection using optimally oriented flux. In *Computer Vision—ECCV 2008: 10th European Conference on Computer Vision, Marseille, France, October 12-18, 2008, Proceedings, Part IV 10*, pages 368–382. Springer, 2008.
- [29] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, and Z. Wang. Photo-realistic single image super-resolution using a generative adversarial network. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

- [30] R Leitgeb, C. Hitzenberger, and A. F. Fercher. Performance of fourier domain vs. time domain optical coherence tomography. *Optics Express*, vol. 11, no. 8, pp. 889-894, 2003.
- [31] Rainer A Leitgeb and Bernhard Baumann. Multimodal optical medical imaging concepts based on optical coherence tomography. *Frontiers in physics*, vol. 6, pp. 114, 2018.
- [32] Ang Li, Jiang You, Congwu Du, and Yingtian Pan. Automated segmentation and quantification of oct angiography for tracking angiogenesis progression. *Biomedical optics express*, 8(12):5604–5616, 2017.
- [33] H. Li, Y. Yang, M. Chang, H. Feng, Z. Xu, Q. Li, and Y. Chen. Srdiff: Single image super-resolution with diffusion probabilistic models. *Neurocomputing*, 2022.
- [34] H. Li and P. Zhang. Spatio-temporal fusion network for video super-resolution. *Proceedings of the International Joint Conference on Neural Networks*, 2021.
- [35] M. Li, Y. Chen, S. Yuan, and Chen. Q. Octa-500. 2019.
- [36] J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, and R.S. Timofte. Image restoration using swin transformer. *Proceedings of the IEEE/CVF international Conference on Computer Vision*, 2021.
- [37] B. Lim, S. Son, H. Kim, S. Nah, and K. Mu Lee. Enhanced deep residual networks for single image super-resolution. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017.
- [38] Singh M, Zvietcovich F, and Larin KV. Introduction to optical coherence elastography: tutorial. *Opt Soc Am A Opt Image Sci Vis*;39(3):418-430. PMID: 35297425; PMCID: PMC10052825., 2022.
- [39] Y. Mei, Y. Fan, and Y. Zhou. Image super-resolution with non-local sparse attention. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021.
- [40] L. Mescheder, M. Oechsle, M. Niemeyer, S. Nowozin, and A Geiger. Occupancy networks: Learning 3d reconstruction in function space. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019.
- [41] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66, 1979.

- [42] J.J. Park, P. Florence, J. Straub, R. Newcombe, and S. Lovegrove. DeepSDF: Learning continuous signed distance functions for shape representation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019.
- [43] Adam Paszke, Abhishek Chaurasia, Sangpil Kim, and Eugenio Culurciello. Enet: A deep neural network architecture for real-time semantic segmentation. *arXiv preprint arXiv:1606.02147*, 2016.
- [44] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.
- [45] S. Peng, M. Niemeyer, L. Mescheder, M. Pollefeys, and Geiger. Convolutional occupancy networks. *Proceedings of the Computer Vision–ECCV*, 2020.
- [46] Lea Querques, Chiara Giuffrè, Federico Corvi, Ilaria Zucchiatti, Adriano Carnevali, Luigi A De Vitis, Giuseppe Querques, and Francesco Bandello. Optical coherence tomography angiography of myopic choroidal neovascularisation. *British Journal of Ophthalmology*, 101(5):609–615, 2017.
- [47] Bhardwaj Rishav, Janarthanam Jothi Balaji, and Vasudevan Lakshminarayanan. Owlslr: Overlapping windows on semi-local region for image super-resolution. *Journal of Imaging* 9, no. 11: 246, 2023.
- [48] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18*, pages 234–241. Springer, 2015.
- [49] Philip J Rosenfeld, Mary K Durbin, Luiz Roisman, Fang Zheng, Andrew Miller, Gillian Robbins, Karen B Schaal, and Giovanni Gregori. Zeiss angioplex™ spectral domain optical coherence tomography angiography: technical aspects. *OCT angiography in retinal and macular diseases*, 56:18–29, 2016.
- [50] S. Saito, Z. Huang, R. Natsume, S. Morishima, and H. Kanazawa, A.; Li. Pifu: Pixel-aligned implicit function for high-resolution clothed human digitization. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019.

- [51] M.S.M. Sajjadi, B. Scholkopf, and M. Hirsch. Enhancenet: Single image super-resolution through automated texture synthesis. *Proceedings of the IEEE International Conference on Computer Vision*, 2017.
- [52] Wasim A Samara, Abtin Shahlaee, Murtaza K Adam, M Ali Khan, Allen Chiang, Joseph I Maguire, Jason Hsu, and Allen C Ho. Quantification of diabetic macular ischemia using optical coherence tomography angiography and its relationship with visual acuity. *Ophthalmology*, 124(2):235–244, 2017.
- [53] W. Shi, C. Jose, H. Ferenc, T. Johannes, A.P. Andrew, B. Rob, R. Daniel, and Z. Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.
- [54] V. Sitzmann, J. Martel, A. Bergman, D. Lindell, and G Wetzstein. Implicit neural representations with periodic activation functions. *Adv. Neural Inf. Process. Syst.*, 2020.
- [55] Bashir S.M.A., Y. Wang, M. Khan, and Y. Niu. A comprehensive review of deep learning based single image super-resolution. *PeerJ Comput. Sci*, 2021.
- [56] Richard F Spaide, James G Fujimoto, Nadia K Waheed, Srinivas R Sadda, and Giovanni Staurengi. Optical coherence tomography angiography. *Progress in retinal and eye research*, 64:1–55, 2018.
- [57] IIBM Spss et al. Ibm spss statistics for windows, version 20.0. *New York: IBM Corp*, 440:394, 2011.
- [58] R. Tan, Y. Yuan, R. Huang, and J. Luo. Video super-resolution with spatial-temporal transformer encoder. *In Proceedings of the IEEE International Conference on Multi-media and Expo (ICME)*, 2022.
- [59] O. Thawakar, P.W. Patil, A. Dudhane, S. Murala, and U. Kulkarni. Image and video super resolution using recurrent generative adversarial network. *Proceedings of the 16th IEEE International Conference on Advanced Video and Signal Based Surveillance AVSS*, 2019.
- [60] VK Viekash, Janarthanam Jothi Balaji, and Vasudevan Lakshminarayanan. Fazseg: a new software for quantification of the foveal avascular zone. *Clinical Ophthalmology*, pages 4817–4827, 2021.

- [61] Nadia K Waheed, Eric M Moulton, James G Fujimoto, and Philip J Rosenfeld. Optical coherence tomography angiography of dry age-related macular degeneration. *OCT Angiography in Retinal and Macular Diseases*, 56:91–100, 2016.
- [62] X. Wang, L. Xie, C. Dong, and Y. Shan. Real-esrgan: Training real-world blind super-resolution with pure synthetic data. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021.
- [63] X. Wang, K. Yu, C. Dong, and C.C. Loy. Recovering realistic texture in image super-resolution by deep spatial feature transform. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018.
- [64] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, C.L. Chen, Y. Qiao, and X. Tang. Esrgan: Enhanced super-resolution generative adversarial networks. *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, 2018.
- [65] Miao Zhang, Thomas S Hwang, J Peter Campbell, Steven T Bailey, David J Wilson, David Huang, and Yali Jia. Projection-resolved optical coherence tomographic angiography. *Biomedical optics express*, 7(3):816–828, 2016.
- [66] Miao Zhang, Jie Wang, Alex D Pechauer, Thomas S Hwang, Simon S Gao, Liang Liu, Li Liu, Steven T Bailey, David J Wilson, David Huang, et al. Advanced image processing for optical coherence tomographic angiography of macular diseases. *Biomedical optics express*, 6(12):4661–4675, 2015.
- [67] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu. Image super-resolution using very deep residual channel attention networks. *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.
- [68] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu. Residual dense network for image super-resolution. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018.
- [69] Zheng, Shu, Bai, Yanru, Xu, Zihao, Liu, Pengfei, Ni, and Guangjian. Optical coherence tomography for three-dimensional imaging in the biomedical field: a review. *Frontiers in Physics*, vol. 9, pp. 744346, 2021.
- [70] Z Zhou, MMR Siddiquee, N Tajbakhsh, and J UNet+ Liang. A nested u-net architecture for medical image segmentation (2018). *arXiv preprint arXiv:1807.10165*.