UNIVERSITY OF SUSSEX SCHOOL OF MATHEMATICS AND PHYSICAL SCIENCES

VeinSense: Deep Learning-Powered Dorsal Vein Recognition

Candidate Number:

260867

Supervisor:

Mr Colin Ashby

Submitted for the degree of Master of Data Science

University of Sussex

August 2023

STATEMENT OF ORIGINALITY

This report is submitted as part requirement for the degree of MSc Data Science at the University of Sussex. This is to certify that to the best of my knowledge, this thesis and the work presented is the product of my own labour and has been generated by me as the result of my own original research except were indicated in the text.

Purvesh Mehta

ABSTRACT

Vein detection plays an integral part of medical procedures like cannulation, blood sample collection and intravenous interventions. Accurate vein detection not only ensures successful procedures but also minimises patient discomfort and healthcare provider stress levels during these procedures. In this report, we discuss easy identification of veins by medical practitioners, improving puncture success rates and patient experience. The primary goal is to establish an efficient vein detection system using deep learning techniques focusing on Dorsal Hand Veins (DHV) recognition as the way forward to streamline healthcare procedures Two methods were utilized to achieve accurate vein segmentation. In the first approach, advanced image processing techniques were leveraged for vein segmentation. Evaluation was performed utilizing Intersection over Union (IoU) and Structural Similarity Image Index (SSIM), opening space for deeper learning techniques to be explored during Method 2. Here, advanced neural networks like ResNet-18, ResNet-34, VGG-16 and VGG-19 were used. Although ResNet and VGG were originally designed for image classification, they have been widely used as the backbone networks for semantic segmentation models as they were trained on millions of images. Our pipeline included data preprocessing, annotation using Label Studio, image resizing, color space conversion and normalization before augmenting with human veins; model performance was assessed through Jacquard Index metrics such as Precision Recall IoU etc. Method 1 resulted in an IoU score of 0.23 which increased exponentially after applying Deep Neural Networks; Method 2 consisted of conducting comparative analyses on various architectures. ResNet-18 boasted exceptional IoU values, attesting to its exceptional accuracy in recognizing DHV. Data augmentation involves applying random transformations like rotation, scaling, and flipping to training images [58]. This creates diverse variations, making the model adaptable to different conditions and prevents overfitting, where the model just memorizes data without understanding it; which was instrumental in improving performance. IoU scores increasing from 0.45 for 50 images over 50 epochs to 0.57. Comparative performance analysis indicated that ResNet-18 (IoU = 0.5745) and ResNet-34 (IoU = 0.5664) outshone VGG-16 and VGG-19 with respect to both loss values and IoU scores. ResNet architectures proved highly suitable for vein recognition. Augmentation proved effective at managing data constraints. Our deep learning-powered vein detection approach showed significant promise, leading to the creation of the Streamlit application that medical practitioners could use easily to identify veins quickly.

Table of Contents

Table of Figures

1. Introduction

Vein detection or prediction plays a vital role in various medical applications, such as cannulation, blood sample collection, and intravenous processes. Accurate and efficient vein detection is crucial for ensuring successful procedures and minimising patient discomfort. The prevalent procedure to detect veins by image sensor involves absorption of a specific wavelength of infrared (IR) light due to deoxygenated blood flowing in the veins [4]. Traditional methods like manual observations and blood flow sensing have limitations in accuracy, success attempts of puncture, patient discomfort, and healthcare provider stress, but recent advancements in infrared imaging-based approaches have shown promise. These approaches use the principle of infrared light absorption by veins detected through IR-sensitive cameras [4]. The absorption of IR by the haemoglobin in blood makes the vein area appear darker under IR illumination, which is suitable for liveness detection [6]. Difficult-to-access veins may be categorized by variations in different body characteristics, such as deep veins due to body fat lowering visibility of peripheral veins, dark skin colour, and thin veins in children and females [4].

This project is motivated by the need for reliable and efficient vein detection systems to detect Dorsal Hand vein (DHV) that can improve the quality of healthcare procedures involving veins. Rather than designing a complete vein detection system from scratch, the

focus was on utilizing existing sets of images and leveraging Deep Learning techniques like neural networks to enhance the accuracy and efficiency of vein detection.

The objective of this project was to develop a deep learning-based system for vein detection. The aim is to create a system that can accurately and reliably detect veins in the first attempt, thereby reducing the need for multiple attempts and minimizing patient discomfort. The major goal of our work is to use the combined image processing and deep neural networks knowledge in implementing DHV recognition to get a comparable or higher accuracies than those reported in the previous literature for hand vein verification. To add a novel idea and make the system accessible to medical professionals, a desktop application is developed using Streamlit. Contradictory to what was mentioned in research proposal that technology will display subcutaneous veins on the screen in real time, due to the timeline restriction on project, now the desktop application will highlight the veins not in real time but show the veins of any dorsal hand image uploaded [1]. It is important to note that the hardware requirements for this implementation of the system are beyond the scope of this research. The focus was on primarily on model training and testing aspects with results of evaluation metrics of the vein detection system and developing a desktop application ready to be used by medical practitioners.

This study includes two methodological phases which are competing approaches. Method-1 utilizes image processing-based DHV detection; while Method-2 utilizes Resnet-18, Resnet-34, VGG-16 and VGG-19 models to detect veins. Method-1 highlights edges along with veins which were refined during Method 2 for pinpoint vein detection. VGG's pre-trained features adapt well to vein detection, while ResNet's residual blocks ease training deep networks without performance degradation. These models enhance feature extraction, pattern capture, and segmentation accuracy. Despite their classification origins, ResNet and VGG are commonly employed as backbone networks in semantic segmentation due to their learned features applicability. Semantic segmentation involves assigning object labels to each pixel using CNN-based feature learning and a decoder network.

Method-1 presents results quantitatively through Structural Similarity Index (SSIM) and Intersection over Union (IoU) [9][10] metrics to provide a complete assessment of image similarity and pixel accuracy, which serves to demonstrate how effective Method-1 was at capturing vein details and their context. IoU (Intersection over Union) is a commonly used metric in semantic segmentation [2] that is beneficial because it allows for the detection of higher-order inconsistencies between ground truth segmentation masks. By comparison, Method-2 employs Resnet and VGG models for deep neural network-based vein detection, measuring their performance via loss, IoU, precision and recall metrics that show strength of their architectures when it comes to vein detection performance. This work achieved higher IoU scores than the papers mentioned in literature section.

2. Literature Review

Vein detection or prediction plays a crucial role in various medical procedures, such as cannulation, blood sample collection, and IV processes. Traditional methods for vein detection have limitations, especially in cases involving hidden or difficult-to-access veins. In recent years, there has been a growing interest in using infrared (IR) and near-infrared (NIR) light for accurate vein detection [4]. Also, use of IR and NIR light allows for better visualization of subcutaneous veins and overcomes challenges related to body characteristics, skin types, and colour. This section will review some papers that explore various aspects of vein detection using NIR.

2.1 Purpose of Literature Review

The purpose of this section is to provide an overview of the existing vein detection or prediction techniques. By reviewing the current state of the field, I aim to identify the limitations that this proposed research will address. DHVs are preferred over palm vein patterns since palm skin is much thicker than the dorsal side [6] [8].

2.2 Classification based on Previous Studies

Existing studies conducted by several researchers including Vito [5] have used Deep neural Networks to train on features while Himani et al (4) uses Image Processing Techniques for DHV recognition. The existing studies by Sougato et al [6] have also explored advanced deep learning architectures such as Siamese and Triplet networks for DHV recognition [6]. The shallow CNN and handcrafted feature methods have shown success in DHV recognition. The possible reasons for that might be the size of the small datasets with an elevated level of uniformity due to the same image-capturing environment and sensors [6]. A multi-modal approach used by Kumar et al. [7] utilizing both finger-vein and fingerprint to perform identification highlighted a superior performance compared to conventional finger-vein identification methods. which may make it harder for the image sensors to capture a superior quality image [6].

Himani et al. [4] conducted a study where they developed an experimental setup utilizing an infrared-sensitive camera and LED arrays with adjustable angles of incidence. They acquired infrared images at various wavelengths and incident angles and applied image enhancement techniques, including adaptive histogram equalization, median blur filtering, and Gaussian blur smoothing. The study implemented an Image Processing-based algorithm for real-time vein detection, involving steps such as grayscale conversion, contrast enhancement, noise removal, and smoothening. The enhanced images successfully allowed for clear visualization of veins on the screen. The findings of the study revealed that the optimal illumination settings for effective vein detection on human hands involved an incidence angle between 0 and 15 degrees to the camera plane. Furthermore, using a combination of multiple wavelengths ranging from 850 nm to 940 nm yielded better visualization of hidden veins compared to using a single wavelength within that range.

Paper by Sougato et al [6] investigated recognition of Dorsal Hand Vein (DHV) as a biometric modality using deep learning techniques. Three CNN models were implemented: a custom six-layer CNN model, a Siamese network using a three-layer CNN model, and a Triplet network utilizing a pre-trained Resnet50 model. The experiments were conducted on a large-scale dataset obtained by combining two publicly available benchmark datasets. Results showed that the shallow custom CNN model outperformed the Siamese network and the Triplet network. The recognition performance is then evaluated using False Accept Rate (FAR) and False Reject Rate (FRR) [6].

Paper by Badawi [59] an outstanding scholarly contribution in the area of DHV recognition stands out amongst existing efforts, by unveiling an ingenious hand vein verification system. By harnessing advanced imaging techniques and computational algorithms, this system examines these intricate vascular networks with great detail; hoping to unlock its potential as an identity verification solution. To evaluate and survey its effectiveness across an expansive and varied dataset comprised of both individuals of differing ages and genders - giving an element of reality by reflecting real world variability in testing scenarios. Performance metrics were Efficiency, Sensitivity, Specificity, False Accept Rate (FAR), False Reject Rate (FRR),

Equal Error Rate (EER), and Receiver Operating Characteristic (ROC), providing a full picture of its operational proficiency. This study contributes to biometric recognition methods through empirical substantiation across diverse datasets combined with analysis of its performance metrics.

2.3 Advantages and Limitations of Current Methods

The existing vein detection methods offer several advantages like providing non-invasive alternatives to traditional venepuncture procedures, reducing patient discomfort. These methods also offer real-time visualization of veins on screen or projecting on patients' regions of interest, aiding healthcare professionals in accurate needle insertion. Furthermore, techniques based on NIR imaging and deep learning algorithms have demonstrated high accuracy in vein detection. However, some current methods also have limitations. They may struggle with detecting hidden or deep veins, as the visibility of veins can be affected by factors like tissue thickness and colour contrast limiting their generalizability. I will try to address and correct this limitation. Ambient light interference can also pose challenges to accurate vein detection.

2.4 Identification of Research Gaps and Opportunities for Improvement

Several research gaps and opportunities for improvement have been identified. Initially, the accuracy and robustness of vein detection methods need enhancement, for challenging cases involving thin or deep veins or presence of root hair. Optimizing the angle of incidence [4] and positioning of the imaging device on the skin can lead to improved vein detection accuracy.

Using advanced deep learning techniques, such as attention mechanisms or generative models, has the possibility of enhancing efficiency and accuracy. Integration of hardware systems with intelligent algorithms, like reinforcement learning, can contribute to achieving precise alignment between the projected image and the detected mask as shown by Vito et al [5].

It is important to provide exact results and use relevant metrics in research papers to evaluate the performance of the models accurately which we have used, which were found to be missing in most of the research papers. Also, incorporating a variety of metrics for assessing model performance is crucial for a comprehensive analysis of the results obtained. By incorporating advanced imaging technologies, novel algorithms, and innovative hardware-software integration approaches, the accuracy and efficiency of vein detection systems can be significantly improved.

3. Dataset

The dataset used in this research consists of two databases of images collected DHV for a biometric recognition project. The purpose of this dataset is to serve as a reference and support research in the field. The databases were collected as part of a larger biometric project conducted in 2007 and 2008 at Pontificia Universidad Javeriana in Bogota, Colombia [12] [13]. The databases contain images of the DHV of participants, captured using a specially designed image acquisition system.

The image acquisition system consisted of a box-like structure with dimensions of 32×29×35 cm (L×W×H), made of wood and coated with extruded polystyrene foam. The bottom of the structure was covered with black foam, and the top had a custom illumination mechanism and a hole for the camera. The system included LED lamps with infrared light emitting diodes (LEDs) for illumination, and a camera (ISG LW-1.3-S-1394) connected to a computer using the IEEE-1394 interface for image capture. An infrared transmitting filter was added to the camera lens to capture images in the infrared spectrum [13].

To enhance image contrast and homogenize colour distribution, a contrast-enhancing control system was implemented. This system controlled the integration time of the camera, affecting the amount of illumination and, consequently, the contrast of the captured image. A control algorithm was developed to determine the contrast value of each frame, and an averaging technique was used to obtain the final captured image.

Two databases: Database 1 and Database 2. Database 1 comprises 138 individuals, with four images per person per hand, resulting in a total of 1,104 images. Database 2 comprises 113 individuals, with three images per person per hand, totalling 678 images. The images have a resolution of 752×560 pixels with 16-bit quantization and are in .tif format. The databases can be accessed online through GitHub repositories [12].

Due to time constraints associated with project's scope, the deep neural network was trained using only part of Database 1 -- specifically 261 images -- for training. As part of this pragmatic approach to meeting project goals within an allotted timeline, these 261 images were manually annotated using Label Studio Cloud Desktop [65] website during this training process ensuring their suitability for model learning. Point to be noted -- The dataset employed for this research differs from that which was utilized by the researchers mentioned in the literature section.

This dataset provides a valuable resource for research in DHV biometrics. It offers a collection of images captured using a specific image acquisition system and includes databases with a considerable number of participants and images. By making these databases publicly available, the authors aim to facilitate further research and exploration in the field of biometric recognition using DHV images [13].

4. Methodology

In this section, we present a detailed methodology that describes two distinct vein segmentation approaches: Advanced Image Processing for Vein Segmentation (Method 1) and Deep Learning Techniques (Method 2). We begin with an overview of key models used in each approach before proceeding into greater depth on their details.

At first, we describe the role of image processing in medical vein segmentation along with an overview of an image processing pipeline. OpenCV and Spyder IDE were utilized, followed by stepwise processing from reading images until final output integration using HSV color space is reached. Evaluation includes Intersection over Union (IoU) and Structural Similarity Image Index (SSIM); sample IoU and SSIM results were summarized along with scatter plots depicting score distributions to provide a solid basis for further exploration into deep learning techniques in Method 2. This creates the backdrop needed for exploring deep learning techniques within Method 2, with further explorations into deep learning techniques.

In Method 2 which is Deep Learning Techniques for Vein Segmentation Advanced neural

networks for automating vein highlighting will be addressed here, beginning with visual representations of input, hidden, and upper layers; then we present four U-Net architectures specifically tailored for vein segmentation (ResNet-34, ResNet-18, VGG-16 and VGG-19). Machine configuration and pipeline flow chart will be introduced. Data preparation involves annotation with Label Studio [65] while image resizing, color space conversion, normalization, and augmentation are among other processes implemented during data processing. Training and evaluation metrics (Jacquard Index, Precision, Recall and IoU) will be discussed, along with comparative performance analyses across architectures as well as model performance discussions. Validation, analysis and result visualization for test images, ground truth data and model predictions round off this section.

4.1 Method 1: Advanced Image Processing for Vein Segmentation

Image processing has emerged as an invaluable asset in medical and scientific research, providing researchers with tools that enable the extraction of intricate details and insights from complex visual data [35,36]. Vein detection relies heavily on its technologies for accuracy and efficiency in locating vein patterns within medical images. Image analysis plays an integral part in improving accuracy and speed when it comes to vein identification and analysis. Utilizing complex algorithms and techniques to uncover information that would otherwise remain unseen by human eyes alone. This has profound ramifications in medical diagnostics and research settings; non-invasively visualizing veins could transform various clinical applications [36]. DHV detection and analysis hold special significance within medical imaging. It offer's valuable biometric characteristics which provide insights for patient identification, cardiovascular health assessment and intervention strategies. Traditional vein detection approaches often fail because of noise, illumination variations and complex hand structures that impede detection. By harnessing image processing techniques for this specific task, we isolate and highlight veins for easier detection. Imagine having access to an accurate and detailed vein viewer that helps doctors and researchers see veins more clearly and precisely while providing medical imaging without needing invasive procedures. Figure 4:1 Below describes the complete Image Processing pipeline.

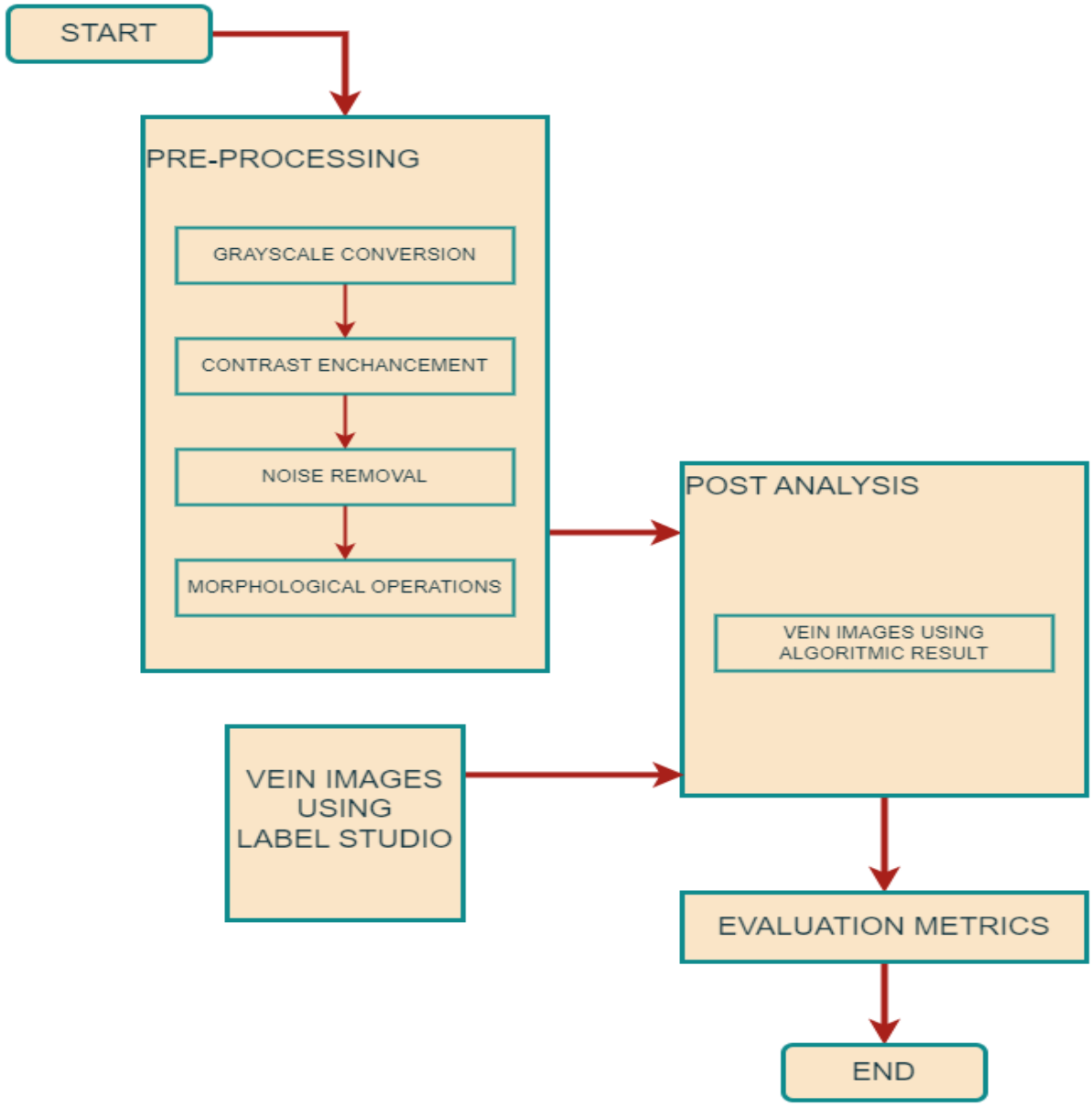

Figure 4:1: Image Processing Pipeline

4.1.1 Description of Code Implementation

OpenCV (Open-Source Computer Vision Library) [14] has become an invaluable tool in computer vision and image processing, providing developers and researchers with a means to explore, manipulate, and gain invaluable insights from visual data [37]. Boasting an open-source philosophy and offering various tools and functions suited for exploration of visual data sets from robots to medical imaging, OpenCV has proven an invaluable asset in computer vision [38]. OpenCV plays an important role in our DHV project, both in terms of accuracy and effectiveness of vein detection/highlighting efforts and in terms of image processing pipeline - it provides an intuitive user interface to handle image acquisition/manipulation/analysis; its extensive functions enable us to enhance DHV visibility by isolating/emphasizing complex patterns more easily.

OpenCV provides everything we need, from reading and preprocessing images to advanced image enhancement techniques [38], for our goal of perfecting vein detection. Leveraging its adaptive thresholding, morphological operations and color manipulation functionalities as a base for accurate vein segmentation and visualization; while its compatibility with various image formats ensures maximum accessibility and versatility for medical research-driven vein analyses.

Spyder IDE creates an advanced image processing pipeline. As an Integrated Development Environment (IDE), Spyder excels at scientific computing and data analysis due to its extensive toolset [25]. User-friendly code editor completes with syntax highlighting and code suggestion helps effortlessly write, test and debug image processing algorithms. IDE's built-in variable explorer greatly facilitates our efficiency when reviewing and manipulating image data, offering greater clarity into image attributes while supporting our vein highlighting techniques [26]. Spyder's image processing tools enables to visualize intermediate outcomes quickly, making fine-tuning of image processing steps much simpler [26]. Together these features have amplified productivity considerably - making Spyder an outstanding choice for this task. Figure 4:2 shows one of the images from dataset.

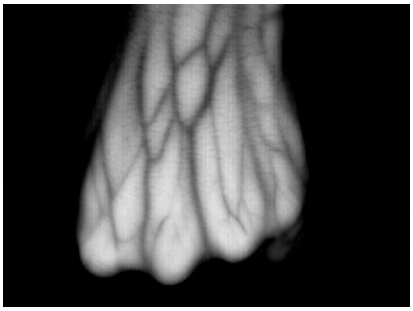

Figure 4:2: Original Image

The first step is to read the input image using the OpenCV library [27]. Loading the image from the specified file path and store it. Grayscale conversion simplifies the image's representation and reduces computational complexity [28]. In the process of developing the pipeline, different kernel size values were worked around used for erosion and dilation to obtain the optimum result for all the DHV images. Different hit-and-trial experiments were conducted to find suitable kernel sizes that effectively removed noise and preserved essential vein structures without decreasing pixel values significantly. Formula for Grayscale Conversion is:

$$Y = 0.299 * R + 0.587 * G + 0.114 * B$$

Where R, G, and B are the red, green, and blue components of the original image, and Y represents the grayscale intensity. Figure 4:3 shows the image after greyscale conversion.

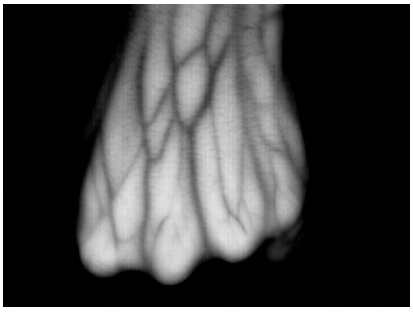

Figure 4:3: Grayscale Output

Next, we utilize adaptive thresholding to create a binary image that emphasizes the veins. Adaptive thresholding is preferred over global thresholding as it accounts for varying illumination conditions in the image [29]. We use cv2.adaptiveThreshold() to apply Gaussian adaptive thresholding to the grayscale image. As part of the optimization process, I had to carefully choose the threshold values to maintain a balance between retaining vein patterns and reducing unwanted noise. Formula for Gaussian Adaptive Thresholding is:

$$T(\mathrm{x},\mathrm{y}) = C - mean(\mathrm{x},\mathrm{y}) + k$$

where T(x, y) is the threshold value at pixel (x, y), C is a constant, mean(x, y) is the local mean intensity around the pixel (x, y), and k is a user-defined parameter to control the threshold. Figure 4:4 gives output after adaptive threshold.

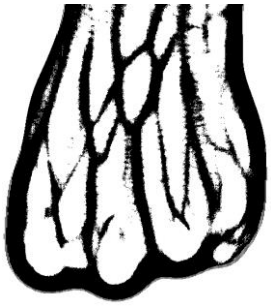

Figure 4:4: Adaptive Threshold

To further enhance the binary image, it is inverted by taking the bitwise not of the thresholded image. This step is essential because adaptive thresholding highlights veins in black against a black background. By inverting the binary image, vein in black foreground and background in white color. Proper handling of border pixels during morphological operations (erosion and dilation) was necessary to avoid either under-removal or over-removal of the vein structures in the image.

$$Inverted\ I\mathrm{mage}(\mathrm{x},\mathrm{y}) = 255 - Binary\ I\mathrm{mage}(\mathrm{x},\mathrm{y})$$

where Inverted Image and Binary Image represent the pixel intensity values of the inverted and binary images. Figure 4:5 illustrates output of the inverted image.

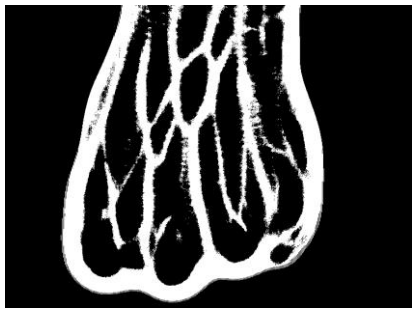

Figure 4:5: Inverted Image

The binary image contains noise and has irregular boundaries, so morphological operations are used to remove noise and smooth the edges of the veins [30]. Morphological operations involve applying structuring elements to accomplish various image processing goals such as smoothing images, identifying object boundaries and reducing noise while distinguishing separate objects. Dilation and erosion stand out among these operations as two pivotal steps; Dilation work to increase object edges by including nearby pixels; on the other hand, erosion reduces these edges by selectively extracting pixels from outer boundaries to

improve object clarity and distinction.These operations alter each pixel systematically until they achieve the desired transformation. After inverting the image, dilation and erosion is performed with multiple experiments with varying iterations and kernel sizes for erosion and dilation to ensure that the resulting vein-highlighted images achieved the best representation of the DHV without sacrificing critical details. Figure 4:6 and 4:7 depict outcomes of morphological operation.

Erosion Formula:

$$Output(x, y) = min_{i}nput(x + i, y + j)$$

Figure 4:6: Morphological Erosion

Dilation Formula:

$$Output(x, y) = maxinput(x + i, y + j)$$

where Output is eroded/dilated image, Input is the binary image, and (i, j) are the coordinates of the structuring element.

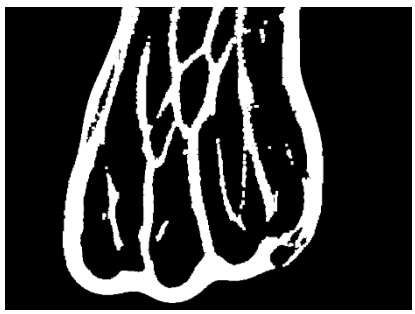

Figure 4:7: Morphological Dilation

Then veins in original image are highlighted. This is done by converting the binary image to an RGB image. This conversion allows to superimpose or mask the highlighted veins on the original image. To visualize the veins clearly, I replaced white pixels in the binary RGB image with red pixels, though different colours can be used as well [31]. This process was carefully tuned to ensure that the veins stand out prominently against the original image without any pixel value degradation. Below figure 4:8 shows RGB image after conversion.

Figure 4:8: Binary to RGB

The highlighted veins are overlaid onto the original image using Gaussian blur for smoothing. Blurring helps to blend the highlighted veins with the original image seamlessly, creating a more natural appearance. The choice of kernel size for Gaussian blur was made after several iterations to ensure that the veins were smoothed without losing important vein patterns.

$$VeinHighlightedImage(x, y) = OriginalImage(x, y) + VeinMask(x, y)$$

To finalize the vein-highlighted image, we convert both the original image and the smoothed vein image to the HSV color space. HSV color space separates color information from brightness, making it easier to manipulate specific color channels. In the HSV color space, we replace the hue and saturation values of the original image with those of the smoothed vein image. Saturation value of the original HSV image is adjusted to ensure that the highlighted veins are visually distinct but still blend well with the original image.

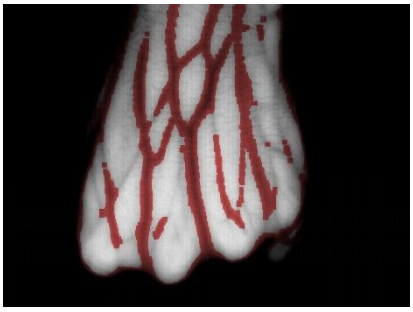

Figure 4:9: Image Masking

The final output in figure 4:9 represents the DHV image with the veins prominently highlighted. This image shows the vein's appearance without significantly altering the original hand image's overall appearance. The implementation was carefully optimized through experimentation and parameter tuning to achieve the best results for various DHV. Each step has novel contribution in highlighting the veins, removing any of the above steps will impact the metric scores.

4.1.2 Evaluation Metrics and Results

IoU Evaluation and Results

IOU (Intersection over Union) is a measure used to quantify the degree of overlap between two bounding regions. A higher IOU value indicates a larger shared area between them and more overlap as indicated by its value [45].

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union}$$

Once model training was complete, evaluating its quality was crucial. Intersection over Union (IoU) emerged as an effective metric for measuring overlap between two bounding boxes or masks; an IoU value of 1 indicates accurate predictions; lower IoU scores indicate less-than-stellar outcomes[20, 46]. Once model training was complete, an IoU metric was employed to assess the extent of overlap in 15 randomly chosen images. The figure 4:10 below illustrates a scatter plot depicting the IoU scores.

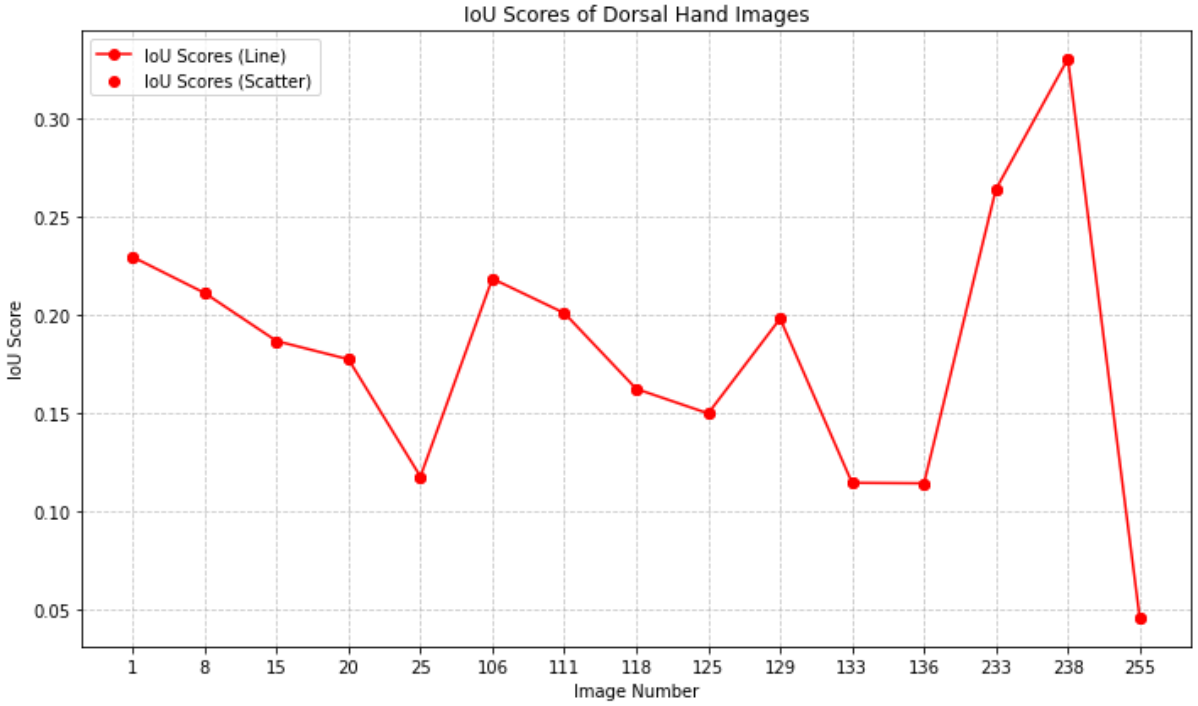

Figure 4:10: IoU Scatter Plot

Structural Similarity Image Index (SSIM)

Structural Similarity Index (SSIM) serves as a vital metric to understand how similar two images are with one another [47, 48, 49, 50]. Going beyond mere pixels-wise comparison, SSIM considers both structural information and perceptual qualities in its analysis - providing an all-inclusive evaluation of how similar two images really are against each other - making SSIM an indispensable asset in tasks such as image comparison, quality evaluation and restoration [47, 48].

SSIM measures how images are perceived by human's visual systems; its components cover luminance, contrast and structure - three key aspects that represent human visual perception - to produce its index value of between -1 to 1. A score of one indicates maximum similarity while one-point dissimilarity exists if values below this are reached.

$$SSIM(x,y) = \frac{(2 \cdot \mu x \cdot \mu y + c_1).(2 \cdot \sigma xy + c_2)}{(\mu_x{}^2 + \mu_y{}^2 + c_1).(\sigma_x{}^2 + \sigma_y{}^2 + c_2)}$$

Where: x and y are the input images being compared. μ_x and μ_y are the mean values while σ_x and σ_y are the standard deviations of x and y. σ_{xy} is the cross-covariance between x and y. c_1 and c_2 are constants. Once model training was complete, an SSIM metric was employed on 15 randomly chosen images along with IoU. The figure 4:11 below illustrates a scatter plot depicting the SSIM scores.

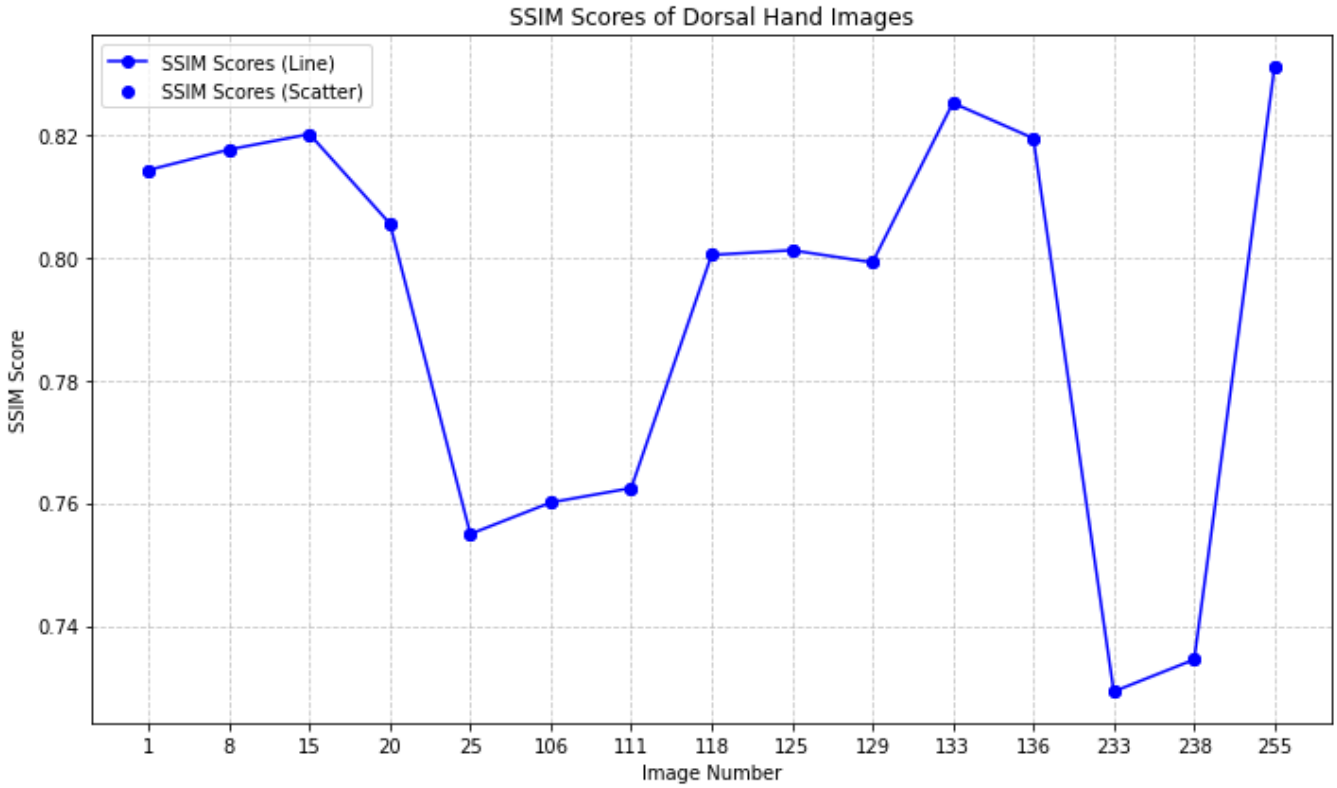

Figure 4:11: SSIM Scatter Plot

Figure 4:12 shows the actual depiction of dataset image along with the ground truth and predicted image by performing Image Processing method.

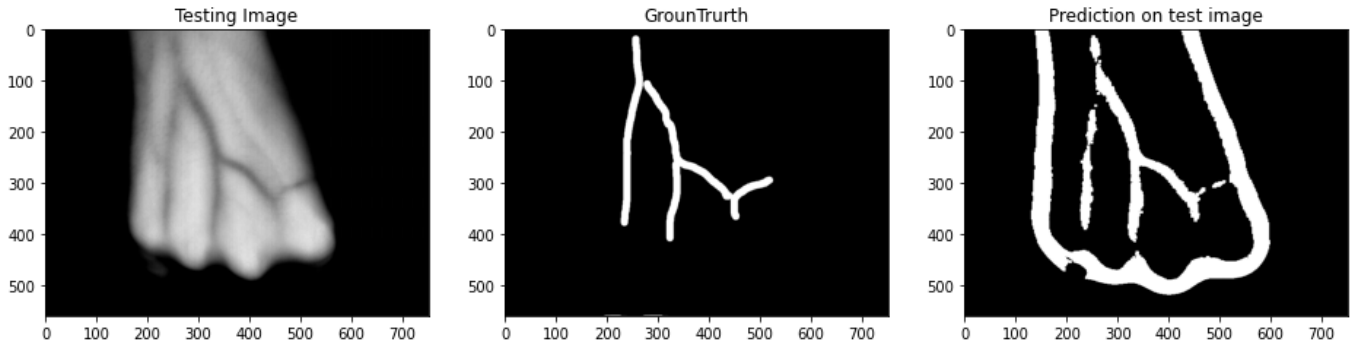

Figure 4:12: Prediction using Image Processing

4.1.3 Reason for using Deep Neural Networks

Image processing methods had one glaring flaw - edges of hands were getting highlighted too. To address these limitations of prior image processing techniques and have another competing approach, method 2 that is deep neural networks will be implemented, improving how veins appear on hand images. Imagine deep neural networks as super smart tools capable of understanding things within pictures; when customized specifically to segmentation tasks they have proven remarkable accuracy at accurately delineating specific structures within images [41,42]. By implementing advanced deep learning methodologies, our aim is to train a neural network using the same dataset of manually annotated DHV images collected through Label Studio Cloud Annotation Website.

Ground truth or gold standard evaluation of models involves providing them with images with veins marked up so they can learn about veins' appearance, then train these tools so they're capable of finding veins in new pictures all by themselves - this should solve our border problem once and for all! Furthermore, this innovative approach promises automatic vein highlighting, thus eliminating challenges encountered through traditional approaches to vein identification.

Deep neural network integration into vein pattern analysis heralds the introduction of an innovative and advanced system capable of providing superior accuracy and insight when it comes to vein highlighting and subsequent analysis.

4.2 Method 2 Deep Learning Techniques

Neural Networks are the functional unit of Deep Learning and are known to mimic the behaviour of the human brain to solve complex data-driven problems [40, 64]. They use representation learning, also known as feature learning, to map input features (analogues to prediction variables in traditional statistics) to an output [33,34, 64]. This mapping process takes place within a series of interconnected layers, each comprising numerous neurons [40]. Figure 4:13 shows general architecture of Neural Networks.

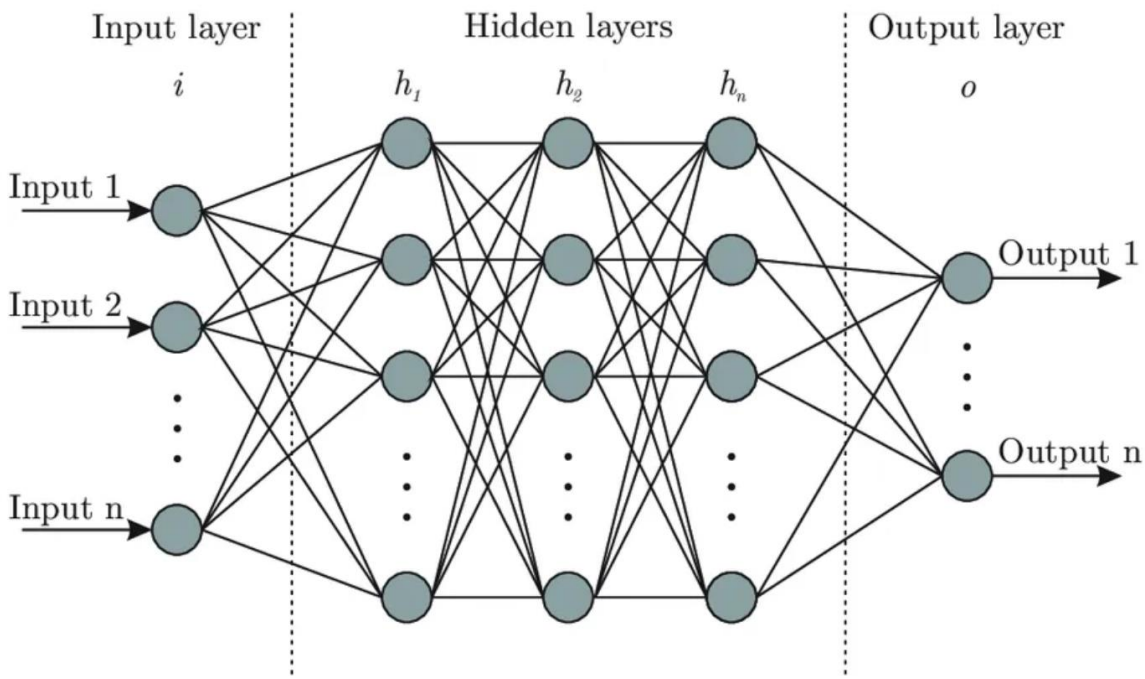

Figure 4:13: Neural Network Architecture

Each neuron serves as a computational unit, and when combined with other neurons, its purpose is to acquire an understanding of all relationship between the input features and the corresponding output [35, 40, 64].

We conducted experiments evaluating four distinct U-Net architectures for vein segmentation tasks.

1) ResNet-34 Encoder U-Net: This model utilizes ResNet-34 encoder, known for its deep architecture and feature extraction abilities [51].

2) ResNet-18 Encoder U-Net: An alternative application utilizes the ResNet-18 encoder, providing an ideal balance of depth and efficiency [51].

3) VGG-16 Encoder U-Net: Utilizing the widely popular VGG-16 encoder, this design leverages its versatility by using VGG for detailed features that need capturing [52].

4) VGG-19 Encoder U-Net: Additionally, we explored the VGG-19 encoder known for its deep architecture and ability to extract richer feature representations [52].

VGGNet, with its VGG-16 and VGG-19 variants, is characterized by its sheer depth and an impressive 138 million training parameters. It achieves this depth with a stack of convolutional layers, max-pooling layers, and fully-connected layers [52, 63]. What distinguishes VGGNet from other architectures is its systematic reduction of input dimensions through each layer, paired with an increment in channel numbers. ResNets architectures have construction of deep networks by stacking residual blocks, enabling networks to reach up to a hundred [51, 63]. ResNet-34 comprises 34 layers with (3x3) convolutional filters, max-pooling layers, and fully-connected layers. Important part being residual connections, which mitigate the vanishing gradient problem and facilitate the training of extremely deep networks [35, 40, 63, 64].

ResNet34 and VGG16 architectures were chosen due to their capabilities that go beyond what was originally designed for image classification. These architectures have widespread application as the backbone networks for semantic segmentation tasks, which complement medical images needed to identify and delineate veins precisely. Semantic segmentation involves labelling objects at pixel level with semantic labels to reveal object boundaries- in our case highlighting veins and learning patterns [63]. Extraction of robust features to identify object boundaries is paramount in semantic segmentation. ResNet34 and VGG16's ability to capture intricate details while their simplicity, which aids training and fine-tuning, demonstrate why these models are suitable for semantic segmentation tasks [51, 52, 63, 64]. While ResNet34 excels at feature extraction, VGG16 offers straightforward architecture. Residual Networks (ResNets) stand out as highly efficient neural network architectures due to their exceptional ability to keep error rates minimal even as network depth increases significantly [63]. Their strategic selection acknowledges that models designed purely for classification can also play an invaluable role when tailored specifically towards dorsal hand vein recognition [63]. Each architecture presents its own distinct approach to vein segmentation within the U-Net framework, using various encoders.

Experiments were conducted using a system equipped with an Intel(R) Core(TM) i7-12700H processor running at 2300 Mhz, featuring 14 cores and 20 logical processors, plus IRIS(R) XE GRAPHICS' 8GB dedicated graphical memory for optimal results.

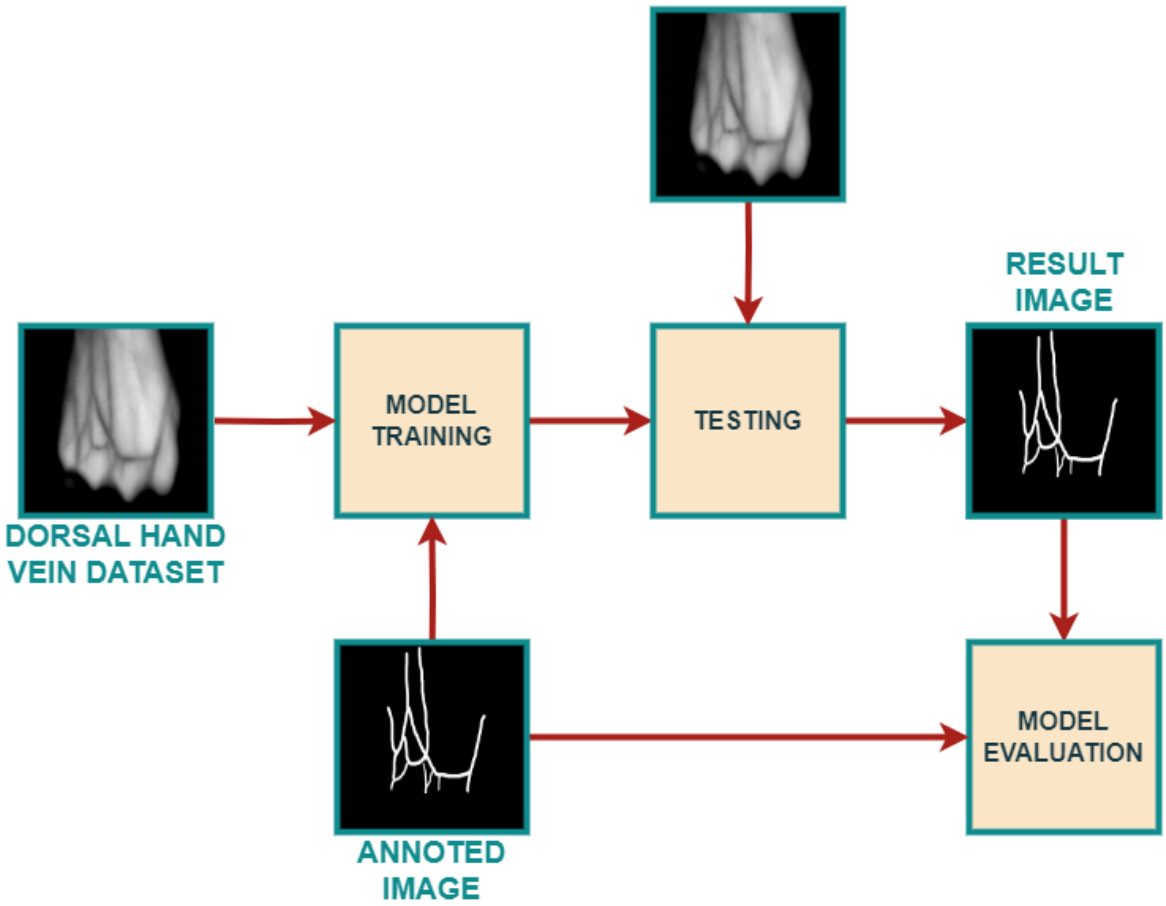

Figure 4:14:Deep Learning Methodology

Figure 4:14 gives a outline of deep learning algorithm for DHV. The research paper we present in-depth investigation of using Deep Learning techniques for vein detection in medical images. Automating this task can greatly increase accuracy and efficiency during various medical procedures; to achieve this we employed Convolutional Neural Networks (CNNs), specifically those featuring U-Net architecture - proven highly efficient at semantic segmentation tasks - for this task. Annotations used as ground truth annotation standards were carried out via Label Studio [65] cloud desktop site annotation service which ensures accurate ground truth annotations as the gold standards when training models after model training was complete.

4.2.1 Data Preparation

Annotations

At this dissertation's data preparation phase, special care was paid to annotation. Of the first dataset's 261 images annotated using Label Studio cloud service manually. With particular attention paid to differentiating between thin and thick veins while considering image variability - an approach adopted so as to ensure accurate and comprehensive annotations which would ultimately strengthen and further augment its robustness for subsequent study and experimentation.

- Original Dataset Image:

The foundation of our research, image 4:15 showcases the raw Dorsal Hand Vein dataset, highlighting the complexity of vein patterns within real-world data.

Figure 4:15: Original Image

- Label Studio Annotation Process

Label Studio was an essential step in the project, streamlining the complex process of annotating Dorsal Hand Vein images quickly and efficiently. Not only was its user-friendly interface impressive; the tool also demonstrated its capacity for complex labeling tasks (notably vein segmentation) with remarkable accuracy and comprehensiveness that greatly helped enhance our dataset preparation phase. With its adaptable handling of diverse data types (including subtle vein structures) enabled by Label Studio we achieved precise annotations which laid a solid basis for future training datasets.

Label Studio proved invaluable in providing seamless interactions between images and annotation, enabling annotators to precisely delineate vein boundaries and collaborative input and revisions, leading to annotations which captured dorsal hand vein nuances accurately. Also annotations done were used for both methods for comparing the efficiency of the models.

Utilizing Label Studio's capabilities, we made sure our training data was of superior quality, leading to the creation of a deep learning model capable of dorsal hand vein recognition. Furthermore, its user-friendly interface combined with complex annotation capabilities made Label Studio an indispensable partner in our pursuit of reliable vein recognition technology. Figure 4:16 shows a snapshot of label studios cloud desktop annotating process.

Figure 4:16: Image Annotation using Label Studio

- Annotated Output from Label Studio:

Label Studio's [65] collaborative annotation and quality control capabilities allowed for us to carefully outline and label veins for Label Studio deep learning models, creating high-quality ground truth data which proved essential in training models to accurately recognize veins under diverse conditions - further increasing research accuracy outcomes. Image in figure 4:17 presents our annotation efforts at work! Our efforts yielded strikingly visible results; note how meticulously veins have been carefully labelled with our annotated dataset.

Figure 4:17: Ground Truth (Annotation)

4.2.2 Data Preprocessing

Before delving into the technical aspects, we discuss the essential data preprocessing steps. The dataset comprises original images in RGB format and their corresponding vein masks in

grayscale. To ensure consistent input sizes for the CNN, all images were resized to a standard dimension of 256 by 256 pixels. Resizing is required to assist uniform feature extraction by the neural network. Also, the RGB images we converted to BGR format and normalized the pixel values in the vein masks to lie within the range [0, 1]. This normalization is important to ease convergence during model training.

To ensure the model's robustness and generalizability, data augmentation techniques were employed during the training process. Data augmentation involves applying random transformations (Figure 4:18) to the training images, such as rotation, translation, scaling, flipping, and shearing [58]. By generating diverse variations of the original images, data augmentation increases the diversity of the training set, helping the model handle various imaging conditions [58]. Additionally, this approach reduces the risk of overfitting, where the model memorizes the training data rather than learning to generalize.

Data Augmentation was deployed due to the relatively small, annotated dataset consisting of only 261 images. By employing various augmentative techniques, effective size was increased so as to permit more varied scenarios to enter into modeling simulation. By mitigating overfitting risk via augmentative approaches and expanding effective size through augmentative techniques, generalization and predictive performance increased dramatically.

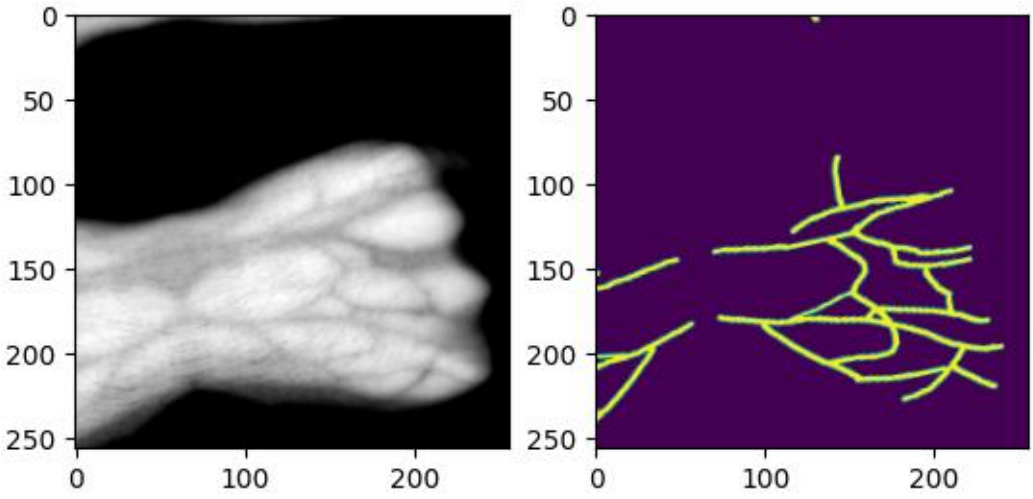

Figure 4:18: Augmentation

The model was trained using the Adam optimizer, a variant of Stochastic Gradient Descent (SGD), which adapts the learning rate based on the gradient information. Each architecture was trained on 50 epochs. The loss function utilized for training was a Jaccard loss.

4.2.3 Evaluation metrics

Evaluating the quality of segmentations and performance of models is an important process in image processing, especially in the medical domain [20]. I will be using various metrics to see how the models are performing. Harnessing the power of accurate vein detection, we aim to maximize true positives (TP), minimize false negatives (FN), reduce false positives (FP), and optimize true negatives (TN).

Jaccard Index Loss Function

The index quantifies the similarity between two finite sample sets, X and Y, using the principle of Intersection over Union (IoU). The IoU is computed as the ratio of the size of the intersection of the sets, |X∩Y|, to the size of their union, |X∪Y|, which is expressed as [43]:

$$IoU = \frac{X \cap Y}{X \cup Y} = \frac{X \cap Y}{X + Y - X \cap Y}$$

As part of segmenting an image, selecting an accurate loss function becomes crucial to its successful processing. A suitable loss function must accurately evaluate each pixel i association by measuring any disparity between its actual ground truth label y_i and current predictions $\hat{y}_i$ [43].

Here, we simplify notations by dropping subscript i for y and $\hat{y}$ for y respectively, simplifying notations by dropping these two symbols from notations altogether.

Rahman and Wang (2016) and Martire et al. (2017) have offered one direct approach for creating loss functions Jd in continuous domain by substituting intersection and union operations with multiplication and addition operations, leading them to formulate Jd accordingly [54, 55, 56, 57]:

$$J_1(y, \hat{y}) = 1 - \frac{(y \cdot \hat{y}) + \epsilon}{(y + \hat{y} - y \cdot \hat{y}) + \epsilon}$$

Cha (2007) [57] presented another variation of J_d by using a normalized inner product with a power of two in its denominator; thus, creating the expression:

$$J_2(y, \hat{y}) = 1 - \frac{(y \cdot \hat{y}) + \epsilon}{(y^2 + \hat{y}^2 - y \cdot \hat{y}) + \epsilon} = \frac{(y - \hat{y})^2}{(y^2 + \hat{y}^2 - y \cdot \hat{y})} + \epsilon$$

This transformation from J_1 to J_2 may be understood within the context of focal loss. Here, the main goal should be to diminish both loss magnitude and gradient for accurate predictions while accentuating them for cases of misclassification [43].

Precision

Precision is the ratio of the correctly +ve labelled by our program to all +ve labelled [21]

$$Precision = \frac{TP}{TP + FP}$$

Recall

Recall, as a metric, measures the ratio of accurate positive predictions among all possible positive predictions. While precision refers to accurate positive predictions within all positive predictions, recall provides insights into positive predictions which were potentially missed; emphasizing missed opportunities.

$$Recall = \frac{TP}{TP + FN}$$

4.2.4 Results and Performance

Comparative Performance of Different Architectures

The table 1 compares various architectural models such as Resnet-18, Resnet-34, VGG-16 and VGG-19 using key evaluation metrics such as loss, Intersection over Union (IoU), precision, recall. Loss values reported for each model ranged between 0.519, 0.5222, 0.5386 and 0.544 respectively, but when considering IoU performance Resnet-18 was found to have the best score with 0.5745 followed by Resnet-34 at 0.5664 while VGG-16 and VGG19 at 0.5573 and 0.5476. Precision values were achieved at 0.8644, 0.8734, 0.8479 and 0.8561 for Resnet-18, Resnet-34, VGG-16 and VGG-19 models respectively; recall data provided insight into their ability to make true positive predictions with recall scores between 0.6869, 0.6648, 0.694 and 6696 for each

Architecture	Resnet-18	Resnet-34	VGG-16	VGG-19
Loss	0.519	0.5222	0.5386	0.544
IoU	0.5745	0.5664	0.5573	0.5476
Precision	0.8644	0.8734	0.8479	0.8561
Recall	0.6869	0.6648	0.694	0.6696

Table 1: Models Performance Comparison

Metrics such as loss, IoU, precision and recall provide insight into a model's predictive abilities across different dimensions. Lower loss values indicate that Resnet-18 and Resnet-34 tend to make more accurate predictions compared to VGG-16 and VGG-19 models. Notably, Resnet-18 outshone its peers when it came to IoU performance; its impressive ability for accurately identifying DHV. Precision values indicate models' capability of making precise positive predictions. Recall scores highlight how effective models are at capturing real positive instances, providing a comprehensive evaluation that guides selection of an architecture best aligned with project objectives, ultimately leading to more informed decision-making for subsequent phases of analysis and development. ResNets excel in deep tasks due to their efficient architecture, maintaining low error rates deep within the network. They shine in tasks like feature extraction, semantic segmentation [63]. As compared with paper by Vito et al [5] IoU and precision of these models are much higher. In paper [5] the IoU maximum score reached 0.454 while we achieved IoU score of 0.5745.

Validation and Performance Analysis:

To assess the model's performance during training and prevent overfitting, we split the dataset into training that is 80% and validation set that is 20%, and separate testing images (15 images). The validation set acts as an independent benchmark, allowing us to monitor the model's generalization capability. During training, we recorded the validation loss and IoU scores to analyze the model's convergence and effectiveness. The file sizes of the trained models after 50 epochs are as follows: ResNet-18 (165 MB), ResNet-34 (281 MB), VGG-16 (272 MB), and VGG-19 (333 MB). By plotting the training and validation loss as well as IoU scores over epochs, we obtained a comprehensive view of the model's training progress in Figure 4:19.

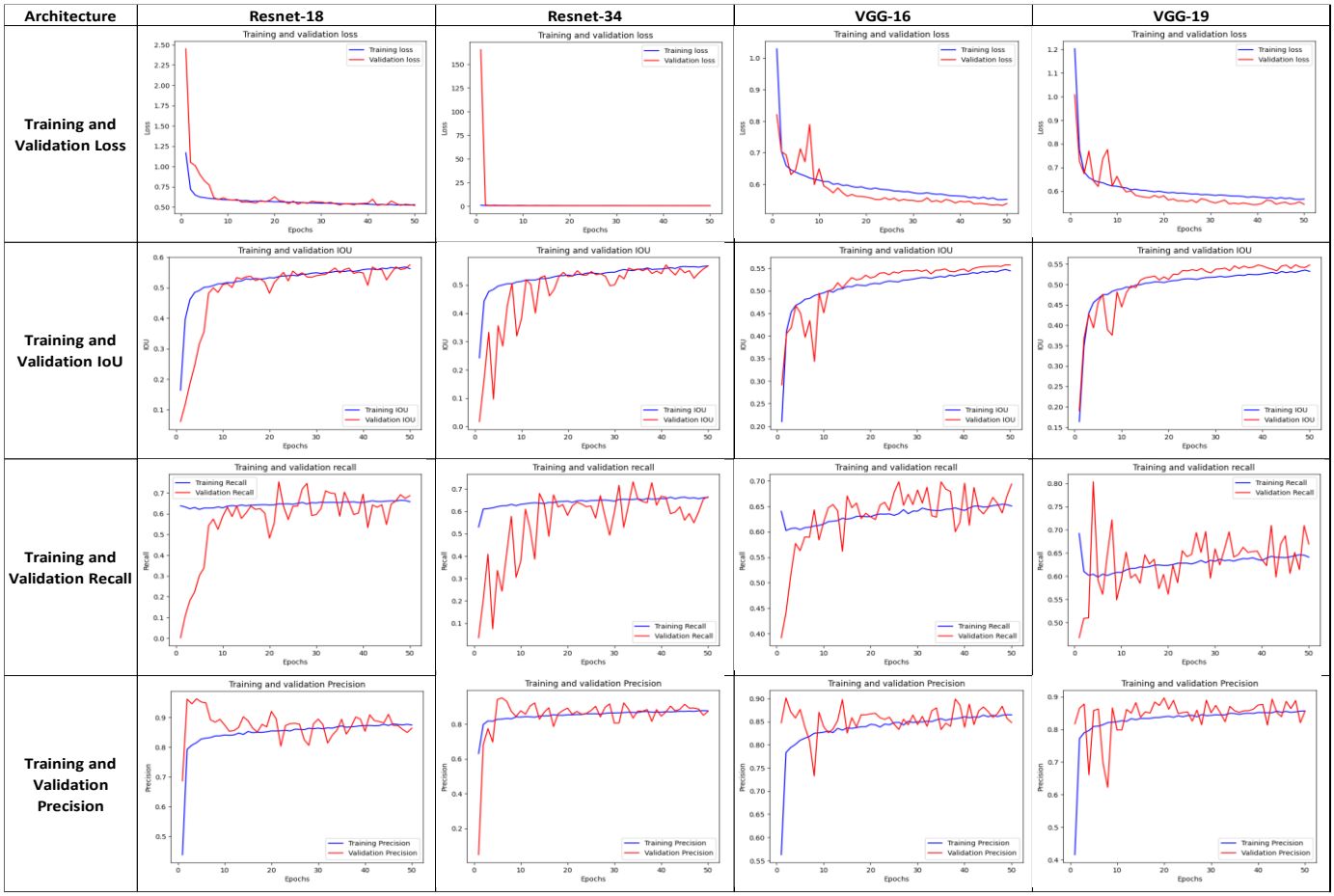

Figure 4:19: Performance of Architectures

Visualization of Results

Following the training phase, the trained model' were saved for subsequent use, and its performance evaluation was conducted on the designated test set images (15 images). As part of an evaluation of the model's efficacy, a visual comparison was carried out (Figure 4:20 and 4:21). This involved side-by-side placement of an original test image, its ground truth vein mask and model-predicted variant (for model prediction). Through visual examination, we were able to assess our model's success in accurately localizing veins as well as determine its degree of alignment between predicted masks and ground truth. Visualizing these results provided us with invaluable insight into our model's strengths and limitations while simultaneously helping to identify areas for enhancement and refinement.

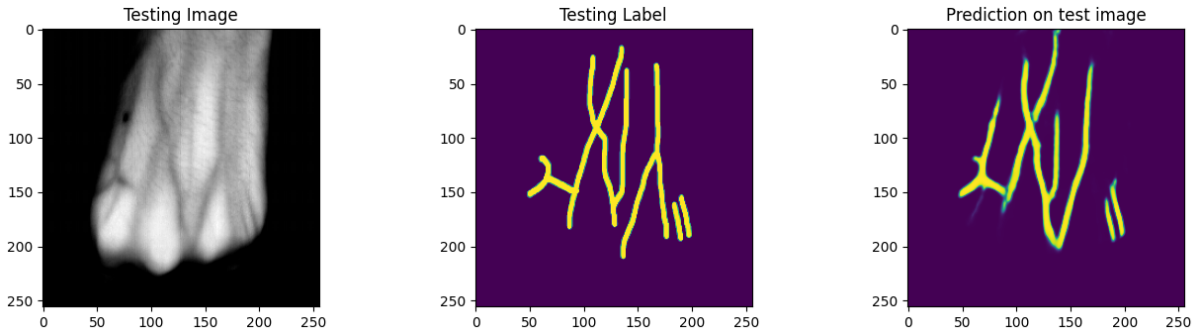

Figure 4:20: Comparative Analysis of Test Image, Ground Truth Vein Mask, and Model-Predicted Variant (Resnet-34)

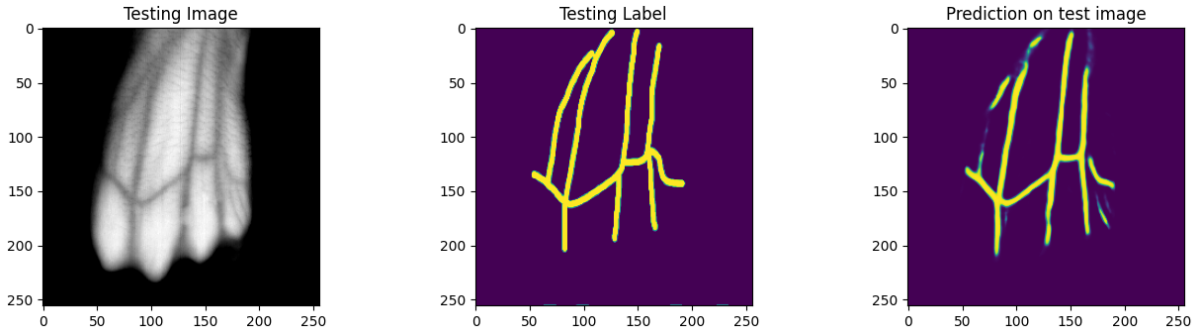

Figure 4:21: Comparative Analysis of Test Image, Ground Truth Vein Mask, and Model-Predicted Variant (VGG-19)

The code utilizes Matplotlib library to generate two side-by-side bar plots to compare training and validation metrics of four architectures (Resnet-18, Resnet-34, VGG-16 and VGG-19). The x axis represents each architecture while the y axis represents metrics scores; bars are color-coded to represent each architecture (blue bars indicate Resnet-18 while orange ones correspond with Resnet-34, grey corresponds to VGG-16 while yellow indicate VGG-19), creating an easy visual comparison among their performance in terms of their output scores during training as well as validation phases of operation to gain insights about architecture performance both when used alongside one another.

Figure 4:22, 4:23 and 4:24 gives an insight into training and validation IoU, Precision and Recall comparison of four models. The 50^{th} epoch results of each model was used to plot the three bar graphs

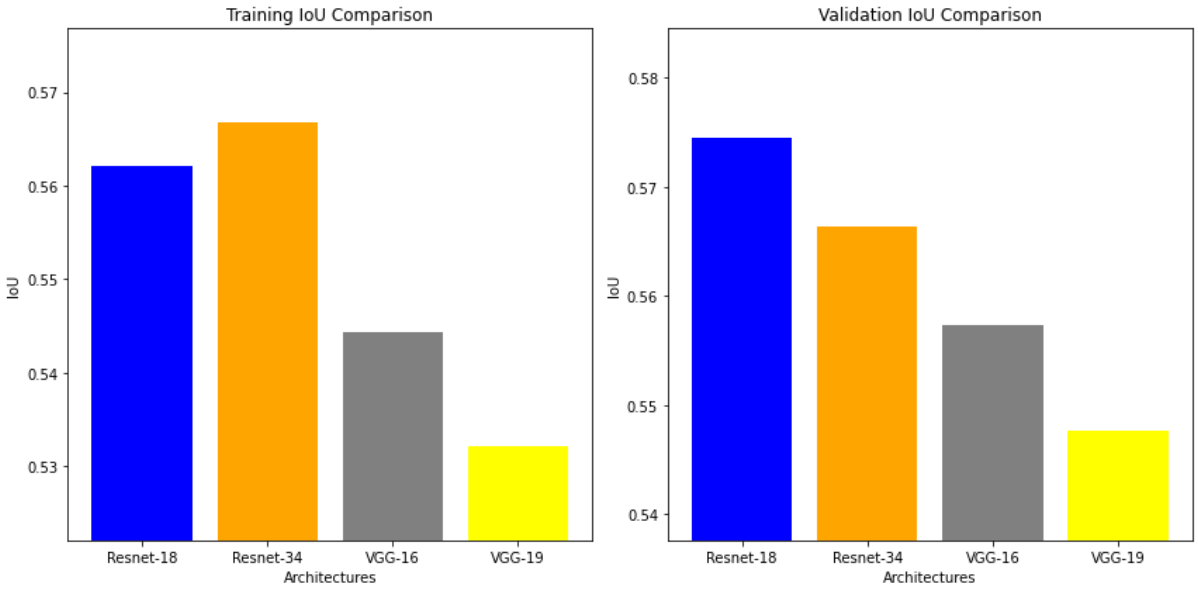

Figure 4:22: IoU Score Comparision

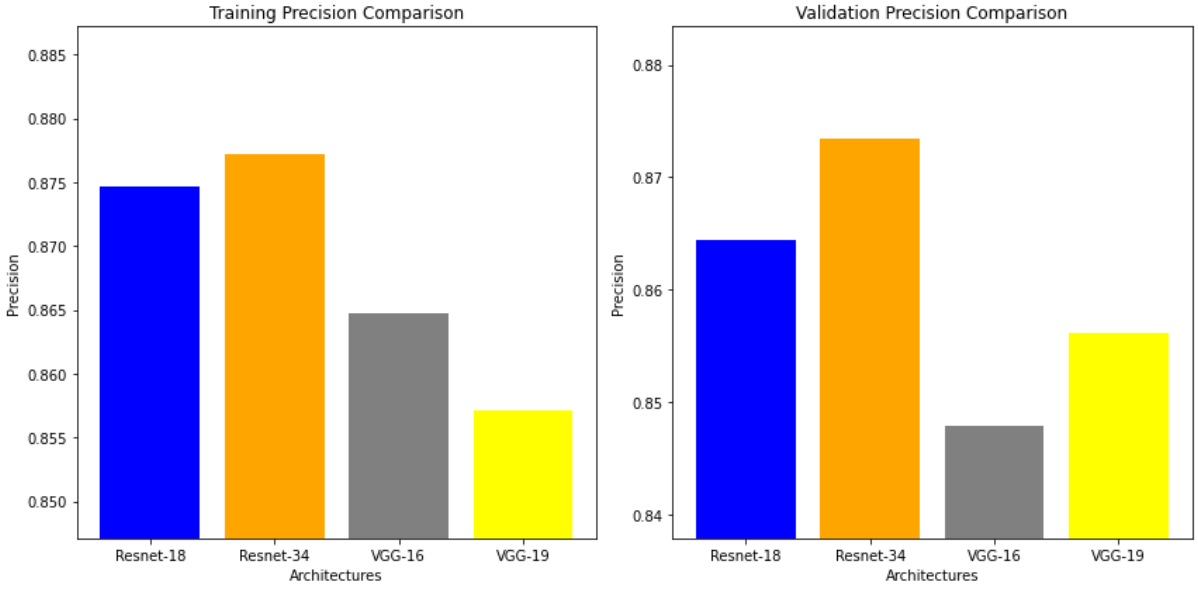

Figure 4:23: Precision Score Comparison

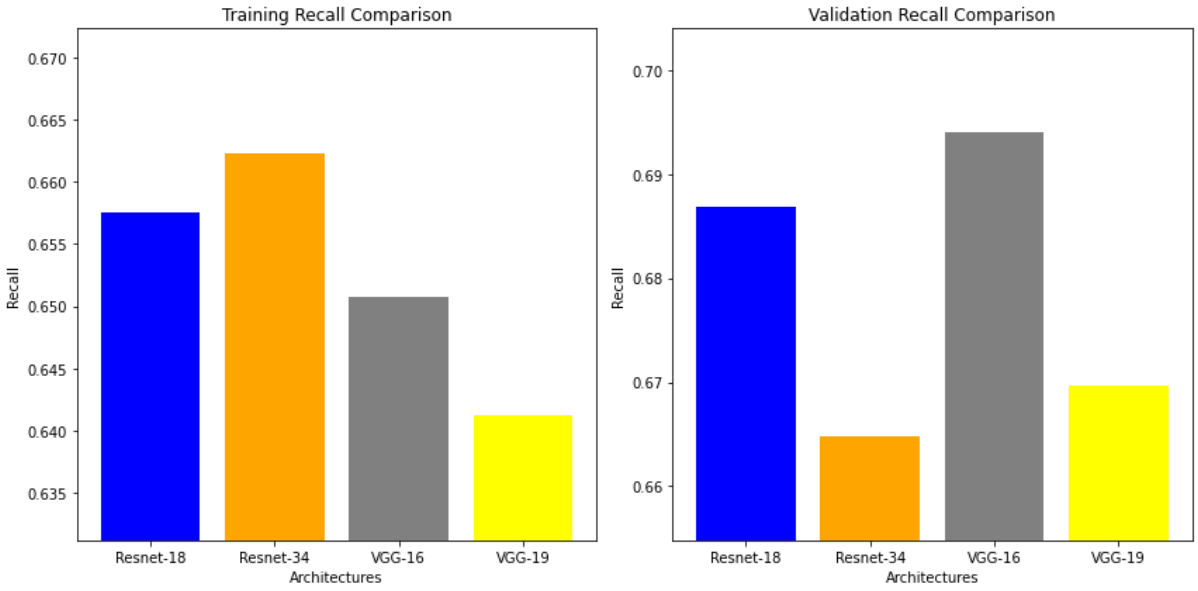

Figure 4:24: Recall Score Comparision

5. Development of a Streamlit Application for Vein Detection and Visualization

5.1 About Streamlit

Streamlit is an open-source library. It has innovative platform that revolutionises data application creation and sharing, integrates seamlessly into project to provide an intuitive yet powerful user experience. Leveraging its capabilities can transform vein detection applications from static tools into user-friendly solutions accessible by medical practitioners, researchers and patients. At the core of Streamlit lies its ease. This project demonstrates its power by taking advantage by crafting an interface which effectively showcases dorsal hand images uploaded and results of image processing/deep learning predictions with just few lines of Python code integrated with the models already trained in four deep neural networks; producing something visually attractive yet functional without most of the complications involved with web development projects.

Real-time updates provided by Streamlit have made DHV more usable. As medical practitioners upload images and apply image processing and deep neural networks techniques, the application automatically updates their results - providing real-time interaction that enhances speed and efficiency when exploring and interpreting data - an integral aspect in the field of medical imaging research and studies.

By harnessing cutting-edge web technologies, Streamlit has allowed this project to provide interactive dashboards, data visualizations and predictions from image processing and deep learning models - while significantly speeding up its development process while maintaining superior user experiences. Furthermore, their extensive documentation and active community have further helped to ease the path in creating user-friendly apps with lasting impacts. Streamlit acts as the driving force of success for any project you undertake, by

helping bridge the gap between complex data processing and user-friendly interfaces, providing faster vein detection solutions in medical images and more streamlined development cycles, automatic updates, and seamless deployment facilitated by it.

In recent years the intersection of machine learning and medical imaging has led to significant improvements in a wide range of healthcare applications. One of these innovations is the development of a new desktop application to aid medical practitioners in vein visualization and detection. This application harnesses image processing techniques and neural networks to enhance analysis of dorsal hands images, providing accurate segmentation and visualisation of veins.

5.2 Methodology

At the core of this Streamlit application, the Streamlit Framework provides an intuitive user interface that allows for the uploading of images and the visualizing of the results. The application has been designed for medical practitioners, to ensure it integrates seamlessly into their clinical workflow. The application interface is characterized by a sidebar containing a project name and a detailed explanation, which outlines the purpose of the application and its potential benefits.

Uploading multiple images is possible in the main application window. This feature, which allows doctors to rapidly analyse a large number of images of the dorsal hands and get insights from them, is vital. This application utilizes two methods to detect and visualize veins: deep neural networks and image processing.

Image Processing - The image processing technique involves a carefully designed sequence of steps that turn the input dorsal images into a visual image highlighting the veins. The application uses adaptive thresholding to accentuate veins against a background. The veins appear in blue on the image, which helps practitioners to analyse and identify vein structures. It also creates downloadable pictures that give a detailed view and enhance documentation. Option has been given so that each image can be downloaded with its two results concatenated as a single image.

Deep Neural networks: This application incorporates machine learning techniques to highlight and predict the veins of dorsal-hand images. The model, which is trained on dataset, uses its learned characteristics to deliver accurate segmentation of veins. Deep learning goes beyond image processing to capture intricate patterns in images eliminating the problem encountered in Method 1 where the side of DHV were also highlighted. The predictions from the neural network appear in red coloured highlighted veins alongside the images.

User Experience Benefits: Figure 5:1 :The design of the application is focused on the user. Images can be uploaded in JPG or PNG format. The progress bar simulates the processing time to give users visual feedback. The interface allows to compare the original image with the results of image processing, as well as neural network predictions which are displayed as a single row with three columns. Images are automatically resized based on the display area of the screen. Option has been given so that each image can be downloaded with its two results concatenated as a single image. Uploaded image size has been restricted to total of 200 MB, to give the best results with least possible time. Also, the size of each uploaded image is displayed. One of the important features being the uploaded and downloaded image has the same image number name when the download concatenated image button is clicked. There is an option to store the images either on the system or on system set-up drive.

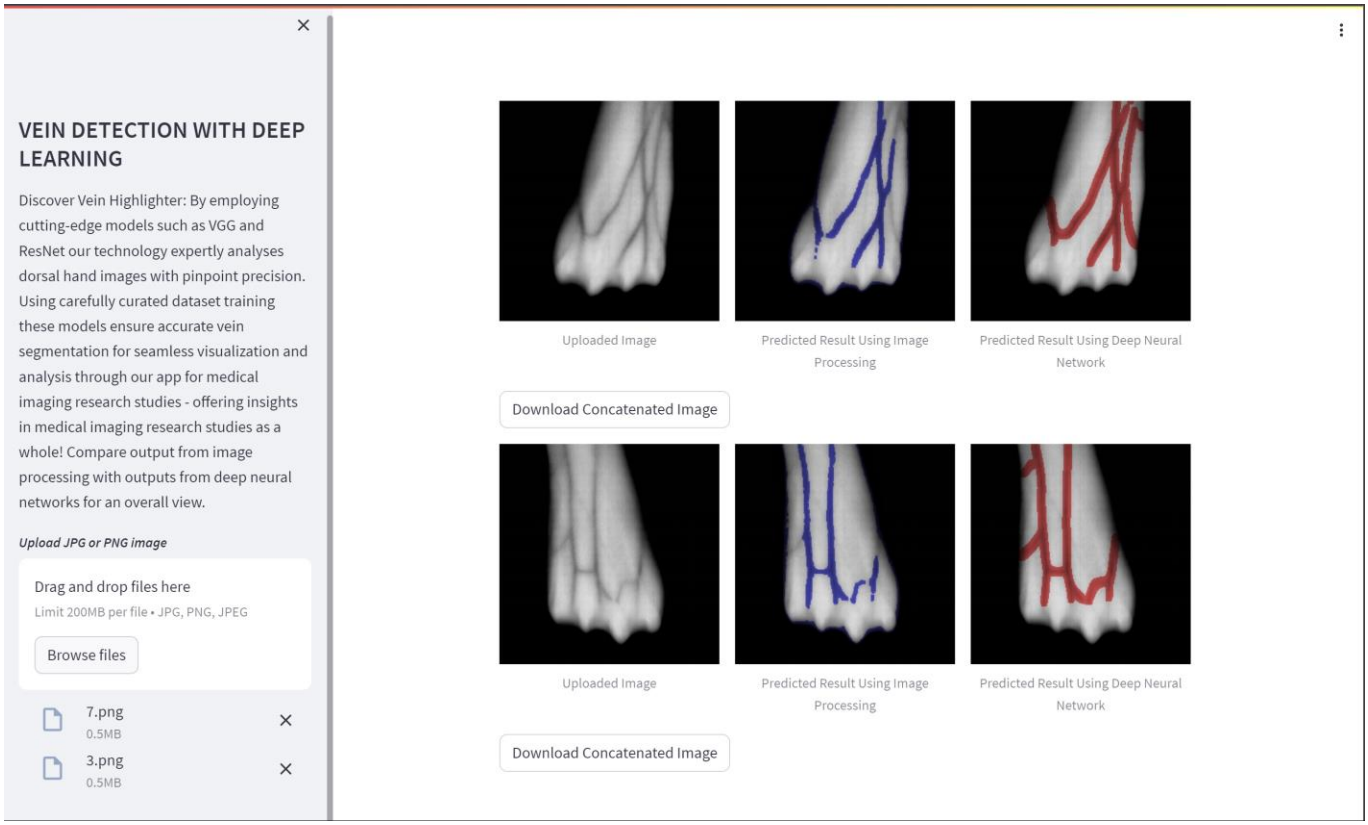

Figure 5:1: A Birds-Eye View of Streamlit Application

Streamlit's application for vein visualization and detection represents a significant improvement in medical image research. The application seamlessly gives two outputs i.e, image processing and deep learning techniques in order to enhance the analysis dorsal-hand images. The application will be a valuable resource for medical practitioners on the field because of its user-friendly design, ability to handle several images at once. Medical imaging is evolving, and innovations such as these can help improve diagnostic accuracy and patient care.

Deploying an application through Streamlit was a challenging task. Seamless deployment process enables anyone to access the application with just a URL address, potentially revolutionizing medical imaging research studies by making accurate vein segmentation and visualization readily available to a larger audience. As our model trained file was bigger than 25 MB (max size limit for uploading a single file on GITHUB) expected we could not deploy it.

6. Conclusion

In conclusion, this dissertation has successfully addressed the vital challenge of dorsal vein detection through the two competing approaches of advanced image processing and deep learning techniques. The significance of accurate vein detection in medical applications cannot be overstated, as it directly impacts patient comfort and healthcare procedures. By leveraging Convolutional Neural Networks (CNNs) particularly ResNet and VGG architectures, this project has demonstrated the potential to revolutionize vein detection. Method 1 employed an image processing pipeline, demonstrating an Intersection over Union (IoU) score of 0.2 and a Structural Similarity Image Index (SSIM) score of 0.79. While effective, this approach also highlighted edges of hand too, which was not required eventually affecting the IoU score. The second method harnessed the power of deep neural networks to enhance vein highlighting accuracy. Residual Network (ResNets) and VGG models, known for their efficiency in segmentation task, were employed to maximize vein detection precision. The project's outcomes have shown that these networks, especially ResNet-18, hold the promise of significantly improving the accuracy of vein detection, as evidenced by superior IoU scores of 0.57 and precision rates which is comparatively higher than the literature papers compared with Vito M [5]. This work not only addresses the need for precise vein segmentation but also bridges the gap between cutting-edge technology and practical medical applications. By harnessing the power of advanced web technologies, we further transformed this research into a user-friendly application powered by Streamlit . This application marks a significant advancement in the practical application of this research. By streamlining the process of vein identification, medical practitioners can now benefit from an efficient and reliable tool for enhanced patient care. This application, when designed to the fullest provides real-time updates, enhancing speed and efficiency in medical imaging research and diagnostic practices. In essence, this project has not only contributed to the field of vein detection but has also laid the groundwork for future innovations that can transform healthcare practices. By leveraging the synergy between image processing, deep learning, and practical application, this research has showcased the potential to make a tangible impact on the medical landscape.

7. Future Work

Future work involves expanding the training dataset to include all 1704 images and exploring the impact on IoU scores, given the notable improvement from 0.45 to 0.57 with increasing image count. Enhancing the streamlit application to allow users to choose from all four models for vein highlighting is another avenue for improvement. Real-time patient data collection for varicose vein disease in clinics could provide valuable insights. Designing an IoT-based system with devices like Raspberry Pi and Nvidia Boards could aid doctors. Integrating a projector as seen in Christie Medical Vein [1] to directly project veins onto the patient's skin presents another exciting possibility.

8. Bibliography

[1] White Paper, Increased Clinical Effectiveness in Pediatric Vascular Access with VeinViewer by Christie Medical Holdings, Inc. [online] Available at: https://veinviewer.com.br/upload/VeinViewer%20reviewed%20by%20parents%20and%20pediatric%20patients.pdf .

[2] Tiu E ., 2020. Metrics to Evaluate your Semantic Segmentation Model. [online] Available at: https://towardsdatascience.com/metrics-to-evaluate-your-semantic-segmentation-model-6bcb99639aa2

[3] Huynh N., 2023. Understanding Evaluation Metrics in Medical Image Segmentation [online] Available at: https://medium.com/mlearning-ai/understanding-evaluation-metrics-in-medical-image-segmentation-d289a373a3f

[4] Himani, M. Kumar, S. Sharma, S. Verma and A. Laddi, "Study and analysis of infrared illumination settings towards detection of hidden veins," 2019 5th International Conference on Signal Processing, Computing and Control (ISPCC), Solan, India, 2019, pp. 171-174, doi: 10.1109/ISPCC48220.2019.8988353.

[5] V. M. Leli, A. Rubashevskii, A. Sarachakov, O. Rogov, and D. V. Dylov, "Near-Infrared-to-Visible Vein Imaging via Convolutional Neural Networks and Reinforcement Learning," 2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV), Shenzhen, China, 2020, pp. 434-441, doi: 10.1109/ICARCV50220.2020.9305503.

[6] S. Bagchi, G. Chanda, A. Agarwal, and N. Ratha, "On Deep Learning for Dorsal Hand Vein Recognition," 2022 IEEE Western New York Image and Signal Processing Workshop (WNYISPW), Rochester, NY, USA, 2022, pp. 1-4, doi: 10.1109/WNYISPW57858.2022.9982726.

[7] Kumar, Ajay, and Yingbo Zhou. ”Human identification using finger images.” IEEE TIP 21.4 (2011): 2228-2244.

[8] Srivastava, S., Bhardwaj, S. and Bhargava, S., 2016. Fusion of palm-phalanges print with palmprint and dorsal hand vein. Applied Soft Computing, 47, pp.12-20.

[9] Padilla, R., Netto, S.L. and Da Silva, E.A., 2020, July. A survey on performance metrics for object-detection algorithms. In 2020 international conference on systems, signals, and image processing (IWSSIP) (pp. 237-242). IEEE.

[10] Taha, A.A. and Hanbury, A., 2015. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC medical imaging, 15(1), pp.1-28.

[11] Tkachenko, Maxim et al., 2020. Label Studio: Data labeling software, [online] Available at: https://github.com/heartexlabs/label-studio.

[12] Wilches-Bernal, F., A database of dorsal hand vein images [online] Available at: https://github.com/wilchesf/dorsalhandveins

[13] Wilches-Bernal, F., Núñez-Álvares, B. and Vizcaya, P., 2020. A database of dorsal hand vein images. arXiv preprint arXiv:2012.05383.

[14] Bradski, G., 2000. The OpenCV Library. Dr. Dobb27;s Journal of Software Tools.

[15] Abadi, Mart&x27;in et al., 2016. Tensorflow: A system for large-scale machine learning. In 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16). pp. 265–283.

[16] Chollet, F. & others, 2015. Keras. [online] Available at: https://github.com/fchollet/keras.

[17] Kluyver, T. et al., 2016. Jupyter Notebooks – a publishing format for reproducible computational workflows. In F. Loizides & B. Schmidt, eds. Positioning and Power in Academic Publishing: Players, Agents, and Agendas. pp. 87–90.

[18] Van Rossum, G. & Drake, F.L., 2009. Python 3 Reference Manual, Scotts Valley, CA: CreateSpace.

[19] Lundh, F., 1999. An introduction to tkinter. [online] Available at: www. pythonware. com/library/tkinter/introduction/index.html

[20] A. A. Taha, A. Hanbury and O. A. J. del Toro, "A formal method for selecting evaluation metrics for image segmentation," 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 2014, pp. 932-936, doi: 10.1109/ICIP.2014.7025187.

[21] Ghoneim G., 2019. Accuracy, Recall, Precision, F-Score & Specificity, which to optimize on? [online] Available at: https://towardsdatascience.com/accuracy-recall-precision-f-score-specificity-which-to-optimize-on-867d3f11124

[22] Varsheni R., 2021. Evaluate Your Model – Metrics for Image Classification and Detection. [online] Available at: https://www.analyticsvidhya.com/blog/2021/06/evaluate-your-model-metrics-for-image-classification-and-detection/

[23] [online] Available at: https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc

[24] IMATEST. SSIM: Structural Similarity Index. [online] Available at: https://www.imatest.com/docs/ssim/

[25] Raybaut, P. (2009). Spyder-documentation. Available Online at: Pythonhosted. Org.

[26] Kadiyala, A., & Kumar, A. (2017). Applications of Python to evaluate environmental data science problems. Environmental Progress & Sustainable Energy, 36(6), 1580-1586.

[27] Bradski, G. (2000). The OpenCV Library. Dr. Dobb27;s Journal of Software Tools.

[28] Kanan C, Cottrell GW. Color-to-grayscale: does the method matter in image recognition? PLoS One. 2012;7(1):e29740. doi: 10.1371/journal.pone.0029740. Epub 2012 Jan 10. PMID: 22253768; PMCID: PMC3254613. [online] Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3254613/

[29] Rosebrock A., 2021.Adaptive thresholding with OpenCV. [online] Available at: https://pyimagesearch.com/2021/05/12/adaptive-thresholding-with-opencv-cv2-adaptivethreshold/

[30] Goyal, M. (2011). Morphological image processing. IJCST, 2(4), 59.

[31] Harris, C.R., Millman, K.J., van der Walt, S.J. et al. Array programming with NumPy. Nature 585, 357–362 (2020). DOI: 10.1038/s41586-020-2649-2.

[32] Ronneberger O, Fischer P, Brox T (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation". arXiv:1505.04597

[33] Goodfellow I, Bengio Y, Courville A, Bengio Y (2016) Deep learning. MIT Press, Cambridge

[34] Miotto R, Wang F, Wang S, Jiang X, Dudley JT (2017) Deep learning for healthcare: review, opportunities and challenges. Brief Bioinform 19(6):1236–1246

[35] Redshift. Deep Learning in Action: Exploring use Cases in Image Recognition and NLP.

[online] Available at: https://www.redshiftrecruiting.com/deep-learning-image-recognition-nlp

[36] R. Deepak Prasanna et al, 2012, Volume 38, Enhancement of Vein Patterns in Hand Image for Biometric and Biomedical Application using Various Image Enhancement Techniques, Procedia Engineering, Pages 1174-1185. Science direct. [online] Available at: https://doi.org/10.1016/j.proeng.2012.06.149

[37] Kauba C, Prommegger B, Uhl A. Combined Fully Contactless Finger and Hand Vein Capturing Device with a Corresponding Dataset. Sensors (Basel). 2019 Nov 17;19(22):5014. doi: 10.3390/s19225014. PMID: 31744197; PMCID: PMC6891606.

[38] Mittal N., August 2023, **Unleashing the Power of OpenCV: A Beginner's Guide to Unlocking Extraordinary Image Processing Techniques!. Medium.** [online] Available at: https://medium.com/@nishitmittal15/unleashing-the-power-of-opencv-a-beginners-guide-to-unlocking-extraordinary-image-processing-903937bc2645

[39] Matson M., 2023.OpenCV: AI Terms Explained. player[zero]. [online] Available at: https://www.playerzero.ai/advanced/ai-terms-explained/opencv-ai-terms-explained

[40] Baheti P, July 2021. The Essential Guide to Neural Network Architectures. v7labs.

[online] Available at: https://www.v7labs.com/blog/neural-network-architectures-guide

[41] Y. Huang, X. Yang, L. Liu, H. Zhou, A. Chang, X. Zhou, R. Chen, J. Yu, J. Chen, C. Chen et al., "Segment anything model for medical images?" arXiv preprint arXiv:2304.14660, 2023.

[42] K. Bera, N. Braman, A. Gupta, V. Velcheti, and A. Madabhushi, "Predicting cancer outcomes with radiomics and artificial intelligence in radiology," Nature Reviews Clinical Oncology, vol. 19, no. 2, pp. 132–146, 2022.

[43] Duque-Arias, D., Velasco-Forero, S., Deschaud, J. E., Goulette, F., Serna, A., Decencière, E., & Marcotegui, B. (2021, February). On power Jaccard losses for semantic segmentation. In VISAPP 2021: 16th International Conference on Computer Vision Theory and Applications.

[44] Shukla L., September 2019. Designing Your Neural Networks. Towards Data Science.

[online] Available at: https://towardsdatascience.com/designing-your-neural-networks-a5e4617027ed

[45] Subramanyam V S, January 2021. IOU (Intersection Over Union). Medium. [online] Available at: https://medium.com/analytics-vidhya/iou-intersection-over-union-705a39e7acef

[46] Sheremet O. July 2020. Intersection over union (IoU) calculation for evaluating an image segmentation model. Towards DataScience. [online] Available at: https://towardsdatascience.com/intersection-over-union-iou-calculation-for-evaluating-an-image-segmentation-model-8b22e2e84686

[47] Li, C., & Bovik, A. C. (2010). Content-partitioned structural similarity index for image quality assessment. Signal Processing: Image Communication, 25(7), 517-526.

[48] Sara, U., Akter, M., & Uddin, M. S. (2019). Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study. Journal of Computer and Communications, 7(3), 8-18.

[49] Channappayya, S. S., Bovik, A. C., & Heath, R. W. (2008). Rate bounds on SSIM index of quantized images. IEEE Transactions on Image Processing, 17(9), 1624-1639.

[50] Hore, A., & Ziou, D. (2010, August). Image quality metrics: PSNR vs. SSIM. In 2010 20th international conference on pattern recognition (pp. 2366-2369). IEEE.

[51] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. CoRR, abs/1512.03385. [online] Available at: https://arxiv.org/abs/1512.03385

[52] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[53] Li, C., & Bovik, A. C. (2010). Content-partitioned structural similarity index for image quality assessment. Signal Processing: Image Communication, 25(7), 517-526.

[54] Deza, M. M. and Deza, E. (2009). Encyclopedia of distances. In Encyclopedia of distances, pages 1–583. Springer.

[55] Rahman, M. A. and Wang, Y. (2016). Optimizing intersection-over-union in deep neural networks for image segmentation. In International symposium on visual computing, pages 234–244. Springer.

[56] Martire, I., da Silva, P., Plastino, A., Fabris, F., and Freitas, A. A. (2017). A novel probabilistic jaccard distance measure for classification of sparse and uncertain data. In de Faria Paiva, E. R., Merschmann, L., and Cerri, R., editors, 5th Brazilian Symposium

[57] Cha, S.-H. (2007). Comprehensive survey on distance/similarity measures between probability density functions. City, 1(2):1.

[58] Brownlee J., April 2019. How to Configure Image Data Augmentation in Keras. Machine Learning Mastery. [online] Available at: https://machinelearningmastery.com/how-to-configure-image-data-augmentation-when-training-deep-learning-neural-networks/

[59] Badawi, A. M. (2006). Hand Vein Biometric Verification Prototype: A Testing Performance and Patterns Similarity. IPCV, 14(3), 9.

[60] Harris, C.R., Millman, K.J., van der Walt, S.J. et al. Array programming with NumPy. Nature 585, 357–362 (2020). DOI: 10.1038/s41586-020-2649-2.

[61] citing J. D. Hunter, "Matplotlib: A 2D Graphics Environment", Computing in Science & Engineering, vol. 9, no. 3, pp.

[62] Van der Walt, S., Sch"onberger, Johannes L, Nunez-Iglesias, J., Boulogne, Franccois, Warner, J. D., Yager, N., … Yu, T. (2014). scikit-image: image processing in Python. PeerJ, 2, e453.

[63] Sharma , V., January 2021. ResNets: Why Do They Perform Better than Classic ConvNets? (Conceptual Analysis). Medium Article. [online] Available at: ResNets: Why Do They Perform Better than Classic ConvNets? (Conceptual Analysis) | by Vanshika Sharma | Towards Data Science

[64] Georgevici, A.I., Terblanche, M. Neural networks and deep learning: a brief introduction. Intensive Care Med 45, 712–714 (2019). https://doi.org/10.1007/s00134-019-05537-w

[65] Tkachenko, Maxim, Malyuk, Mikhail, Holmanyuk, Andrey, & Liubimov, Nikolai. (2020). Label Studio: Data labeling software. Retrieved from https://github.com/heartexlabs/label-studio