Advanced Image Classification for Varicose Veins Detection Using ResNeXt

A Project Report Submitted in the
Partial Fulfillment of the Requirements
for the Award of the Degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING (AI&ML)

Submitted by

CHAKALI DIVYA	21881A6616
SABAVAT VINOD NAYAK	21881A6655
VADLA ROHITH	22885A6607
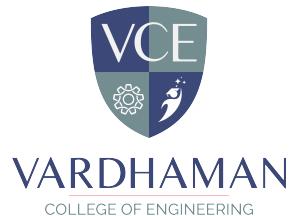
SUPERVISOR
Dr M A JABBAR
Professor and Head
Dept of CSE(AI&ML)

Department of Computer Science and Engineering(AI&ML)

April, 2025

VCE

VARDHAMAN
COLLEGE OF ENGINEERING

Department of Computer Science and Engineering (AI&ML)

CERTIFICATE

This is to certify that the project titled **Advanced Image Classification for Varicose Veins Detection Using ResNeXt** is carried out by

CHAKALI DIVYA	**21881A6616**
SABAVAT VINOD NAYAK	**21881A6655**
VADLA ROHITH	**22885A6607**

in partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science and Engineering (AI&ML)** during the year 2024-25.

Signature of Supervisor   **Signature of the HOD**
Dr. M.A. Jabbar   **Dr. M.A. Jabbar**
Professor and Head   **Professor and Head**
Dept of CSE(AI&ML)   **Dept of CSE(AI&ML)**

Project Viva-Voce held on _____

Examiner

Acknowledgements

Abstract

Varicose veins and chronic venous disease (CVD) are prevalent vascular conditions with significant clinical and quality of-life impacts. Accurate and timely diagnosis of these conditions is crucial for effective management and treatment. This research explores the application of ResNeXt, deep learning architecture, for image classification of varicose veins and staging of varicose veins. ResNeXt's unique split-transform-merge strategy and en hanced feature extraction capabilities offer a promising approach for distinguishing between various venous pathologies from image data. We employ ResNeXt to classify images obtained from venous and other imaging modalities, aiming to improve diag nostic accuracy and assist healthcare professionals in identifying and assessing venous disorders. Our results demonstrate that ResNeXt achieves high classification performance, outperforming traditional methods and contributing to more precise and efficient diagnosis of varicose veins and CVD. This approach highlights the potential of advanced convolutional neural networks in enhancing medical imaging analysis and patient care in vascular health.

Keywords: Varicose veins, ResNeXt convolutional neural networks, CVI, CNN, CEAP, CVD.

Table of Contents

iii

List of Tables

List of Figures

Abbreviations

Abbreviation	Description
ResNet	Residual Networks
ResNeXt	Residual + Next-level Architecture with Aggregated Transformations
CEAP	Clinical,Etiological,Anatomical,Pathophysiological
MRI	Magnetic Resonance Imaging
CNN	Convolutional Neural Network
CVD	chronic venous disease
VGG	Visual Geometry Group

CHAPTER 1

Introduction

1.1 Introduction

Varicose veins represent a common vascular complaint that manifests with dilated, tortuous veins, most often in the legs due to poor blood flow. This disorder develops when the venous valves do not work appropriately, allowing blood to pool and venous pressure to increase. Left untreated, varicose veins will lead to chronic venous disease (CVD) and can create a noticeable physical complaint such as pain, swelling, skin discoloration, ulceration, and in extreme cases, deep vein thrombosis. Therefore, early identification and classification of the venous disorders is needful to allow for effective medical management and treatment. Varicose veins are usually
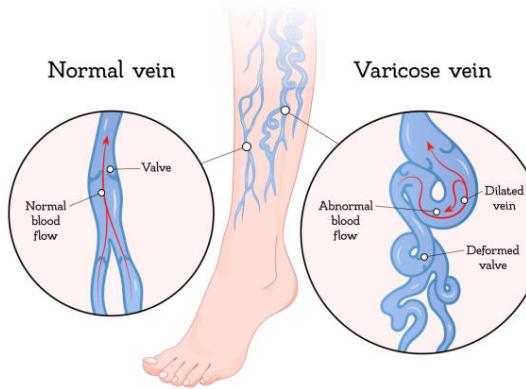
Figure 1.1: Normal and Varicose vein[12]

diagnosed based on physical examination, history of the patient, and Doppler ultrasound examination. Each of these modalities has limitations, including reliance on a physician's experience, subjective diagnoses, and time required for the examination. With modelling and other advances in artificial intelligence and deep learning, automated image classification systems have the potential to improve diagnostic accuracy, reduce human error, and speed the process of diagnosis.

This project presents a superior image classification model based on ResNeXt, a deep CNN that has demonstrated excellent performance, particularly in medical image analysis. ResNeXt was designed with grouped convolutions which provide better means of feature extraction and classification accuracy. Using an extensive dataset of varicose vein images, this research aims to develop an AI model that detects varicose veins and classifies them into variations of severity.

This project includes many important components including a dataset collection, image preprocessing, implementation of a deep learning model, model training, and performance assessment. The proposed method's effectiveness will be validated using a range of important performance metrics, including accuracy, precision, recall, and F1-score. The performance of the proposed method will also be presented alongside the performance of other deep learning architectures to assess the performance superiority of the proposed method in the classification of varicose veins.

This study adds to the expanding domain of healthcare interventions utilizing artificial intelligence in medical diagnostics and enhances the speed, accuracy, and cost-effectiveness of disease identification. This study's results may help healthcare workers make decisions based on credible evidence, enhance patient care, and improve automated diagnostic approaches for vascular diseases.

1.2 Background and Motivation

1.2.1 Background

Varicose veins are a common vascular condition that affects millions of people around the world, particularly frailer older persons and pregnant individuals and, but also people whose jobs require standing for long periods. The condition is caused by venous insufficiency where a valve that is weak or damaged prevents blood from flowing properly and causes veins to enlarge, coil and become visible under the skin. Chronic venous diseases (CVD) associated with varicose veins can have severe complications such as leg ulcers, venous

thrombosis, and skin infections.

Medical imaging plays a crucial role in diagnosing varicose veins, with Doppler ultrasound being the primary diagnostic tool. However, ultrasound imaging requires specialized equipment and trained professionals to interpret the results accurately. This dependence on manual assessment introduces variability in diagnosis and delays early detection, which is critical for effective treatment. As a result of the development of artificial intelligence (AI)
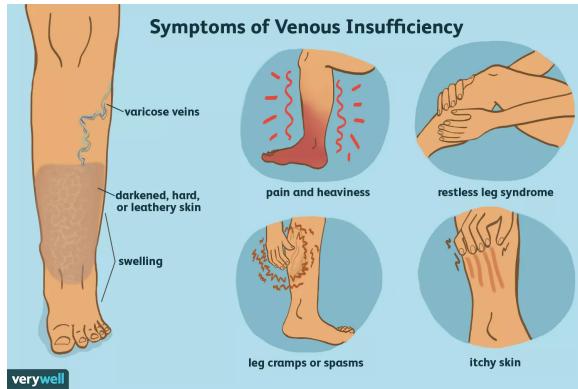
Figure 1.2: Chronic Venous Insufficiency[14]

and machine learning (ML), computer vision techniques have become commonplace in medical diagnostics. Convolutional neural networks (CNNs) have been exceptionally successful in studying medical imagery, like X-rays, MRIs, and ultrasounds, facilitating highly accurate and automated disease detection. However, the application of deep learning within identification of varicose veins is still in its early stages and further research is needed to produce important, usable models that provide support to medical professionals interpreting and classifying varicose veins.

1.2.2 Motivation

The reason for this project is an increasing demand for precision and efficiency in automated diagnostic systems within the health domain. The conventional approaches to detect and classify varicose veins can take a long time to achieve, result in human error, and tend to rely upon excessive levels of medical expertise. An image classification model powered by AI would provide a faster, more accurate, and economical alternative to an evaluation

by a human.

This project builds upon the ResNeXt structure; a deep learning model recognized for its superior accuracy of execution in the classification of medical images. ResNeXt uses varied, grouped convolutions, unlike conventional CNNS, to improve the extraction of features and proves to be a reasonable choice when identifying complex patterns found in images of varicose veins. The development of a greater classification system is aimed at addressing the disparity between the traditional way a clinician would provide a medical diagnosis and AI-led healthcare provision.

The significance of this research lies in its potential to:

- Enhance early detection and treatment of varicose veins, reducing the risk of complications.

- Provide an AI-driven decision-support tool for medical professionals, improving diagnostic efficiency.

- Contribute to the growing body of research on AI applications in vascular disease detection.

- Reduce dependency on specialized medical expertise, making varicose vein detection more accessible in resource-limited healthcare settings.

1.3   Problem Statement

Varicose veins are an important medical ailment affecting millions of people worldwide. The disease arises due to venous insufficiency, with damaged valves in the vein allowing blood to flow backward, which leads to vein dilation, pain, and possible complications, such as ulcers and deep vein thrombosis. Therefore, early detection of the disease is important for preventing a serious medical event; however, current detection methods primarily rely on clinical examination and Doppler ultrasound imaging, which can require expertise and can be time consuming.

Despite advancements in medical imaging, there is still a lack of automated and accurate varicose vein detection systems. Traditional methods of diagnosis

present several challenges:

- **Subjectivity in Diagnosis:** Manual assessment of varicose veins depends on the experience and skill of healthcare professionals, leading to variability in diagnosis.

- **Time-Consuming Procedures:** Conventional imaging techniques, such as Doppler ultrasound, require significant time for analysis and interpretation.

- **Limited Accessibility:** In many regions, access to specialized medical professionals and imaging technology is limited, leading to delayed diagnosis and treatment.

- **Need for Automated Classification:** Existing diagnostic approaches primarily focus on detecting the presence of varicose veins rather than classifying their severity, which is essential for determining appropriate treatment strategies.

The objective of this project is to overcome these hurdles by developing an enhanced image classification system based deep learning for varicose vein detection based on the ResNeXt architecture. The proposed algorithm will produce an automated detection system that will improve diagnostic accuracy and provide a classification of varicose veins by severity. Through AI-based solutions, this research will attempt to close the gap between traditional diagnosis and automated modern health technologies to promote earlier detection and patient management.

1.4 Objectives of the Project Work

The primary goal of this research is to develop an advanced image classification system for varicose vein detection using the ResNeXt deep learning architecture. The objectives of this study are set out below:

- To develop a deep learning-based image classification model capable of detecting varicose veins from medical images with high accuracy.

- To design and implement an automated system that can classify varicose veins into different severity stages, aiding in early diagnosis and treatment planning.

- To preprocess and augment the dataset to enhance model performance and improve the generalization of the classification model.

- To analyze the impact of different hyperparameters, including learning rate, batch size, and number of layers, on model accuracy and efficiency.

- To evaluate the model's performance using standard classification metrics such as accuracy, precision, recall, F1-score, and AUC-ROC.

- To compare the ResNeXt model's effectiveness with other deep learning architectures such as ResNet, VGG, and EfficientNet to establish its superiority.

- To integrate the trained model into a user-friendly interface, making it accessible for medical professionals for real-time varicose vein detection.

- To explore the feasibility of deploying the model in clinical settings for assisting doctors in automated diagnosis and decision-making.

1.5 Organization of the Report

This report is structured into six chapters, each detailing different aspects of the research work. The organization is as follows:

- **Chapter 1: Introduction** – This chapter provides an overview of the research, including the background, motivation, problem statement, objectives, and the structure of the report. It establishes the foundation for the study and highlights the importance of automated varicose vein detection.

- **Chapter 2: Literature Review** – This section presents an extensive review of existing research related to varicose vein detection, deep

learning applications in medical imaging, and the effectiveness of different convolutional neural networks. It critically examines prior studies, identifying gaps and challenges that this research aims to address.

- **Chapter 3: Methodology** – This chapter outlines the technical approach used in the study, including dataset collection, preprocessing techniques, model selection, and implementation of the ResNeXt architecture. It also details the hyperparameters, training process, and optimization strategies applied to improve the model's performance.

- **Chapter 4: Implementation of the Proposed System** – This section describes the step-by-step execution of the proposed deep learning model, including training, validation, and testing procedures. It also discusses how the classification system was fine-tuned for optimal accuracy and efficiency.

- **Chapter 5: Results and Discussion** – This chapter presents the evaluation of the trained model based on various performance metrics such as accuracy, precision, recall, F1-score, and AUC-ROC. It compares the results with other existing models and provides an analysis of the strengths and limitations of the proposed system.

- **Chapter 6: Conclusion and Future Work** – The final chapter summarizes the key findings of the research, emphasizing its contributions to medical imaging and AI-driven healthcare solutions. It also discusses potential improvements and future research directions to further enhance the accuracy and applicability of the model.

CHAPTER 2

Literature Survey

2.1 Introduction

The objective of this chapter is to examine prior studies and studies in the literature pertaining to the detection of varicose veins and deep learning applications in medical imaging. An important part of the review process is surveying the literature to provide an understanding of the current state of the art, where the gaps are in the methodologies, and to provide a basis or context for the proposed work. This chapter will also provide a valuable look at the various machine learning techniques, image processing techniques, and classification models that have been utilized in prior studies to assist in medical diagnosis.

The literature review has been completed with usage of credible sources, including scientific journals, conference proceedings, and technical reports from artificial intelligence, deep learning, and medical imaging domains. Research papers were scrutinized based on methodologies, datasets, performance measures, and findings. The literature was restricted to works that discuss convolutional neural networks (CNN), transfer learning, and deep learning architectures such as ResNet, ResNeXt, and EfficientNet for medical image classification. The purpose of this chapter is to investigate the existing literature and technologies with respect to the detection of varicose veins and deep learning in medical imaging. In order to comprehensively understand the methodologies and the relative strengths and limitations of the prospective studies within the existing literature, an exploration of the feasibility of promoting further research was deemed necessary. This chapter illustrates both traditional and machine learning-based approaches to diagnosing varicose veins, as well as their merits and demerits.

The chapter opens with a review of standard techniques for detecting

varicose veins, which consist of ambulatory venous pressure measurements, clinical evaluation and imaging techniques (Doppler ultrasound and venography). These standard methods can entail significant medical experience and have considerable diagnostic variability. No way. The chapter subsequently explores current artificial intelligence methods, namely deep learning methods that are changing medical image classification. The chapter emphasizes convolutional neural networks (CNNs) in deep learning that can accurately interpret medical images and detect vascular anomalies.

In addition, an analysis of research on AI-based detection of varicose veins is shared, reviewing the various machine learning models, datasets, and performance metrics used in previous studies. This section emphasizes the promise of deep learning models, specifically ResNet and ResNeXt, for enhancing the accuracy of detection and increasing automation. Even with many advancements, there are still clear research voids with concern to lack of data, model generalizability, and computational difficulty. These research voids are explored in detail resulting in a need for more robust, efficient and scalable models to determine if varicose veins are present. At the conclusion of this chapter, a summary of the main findings from the literature will be presented with a view to indicating their relation to the discovered knowledge. The collective findings from the literature are reflective of a body of knowledge informing the proposed work which is designed to build an advanced deep learning model based on the ResNeXt architecture to automatically and accurately detect advisory varicose veins.

2.2 Review of Prior Research

The detection and classification of varicose veins using deep learning have gained significant attention in recent years due to the advancements in artificial intelligence (AI) and medical imaging. Traditional methods for diagnosing varicose veins primarily rely on clinical evaluations, imaging techniques, and manual assessment by medical professionals. However, these conventional techniques often present challenges in terms of accuracy, efficiency, and accessibility. The rise of deep learning models, particularly convolutional neural networks

(CNNs), has introduced new possibilities for automating varicose vein detection, improving diagnostic precision, and facilitating early intervention.

This presents a comprehensive review of previous research in varicose vein detection, analyzing methodologies, techniques, and findings. The existing literature is categorized into traditional clinical approaches, deep learning-based classification, and optimization strategies for improved accuracy. Additionally, this section identifies key research gaps and areas requiring further investigation.

2.2.1   Traditional Approaches for Varicose Vein Detection

Historically, varicose vein diagnosis has been conducted through physical examinations, Doppler ultrasound imaging, and venography. These methods allow healthcare professionals to assess vein abnormalities, identify reflux patterns, and measure blood flow efficiency. Rabe et al. [4] conducted an epidemiological study on chronic venous disorders across diverse populations, emphasizing the importance of early diagnosis through comprehensive clinical evaluations. Similarly, Borde and Savrasov [6] explored mathematical modeling techniques for ultrasound heating to assist in varicose vein diagnosis.

While these traditional methods remain essential for clinical practice, they present limitations such as dependency on operator expertise, high inter-observer variability, and time-intensive procedures. Furthermore, manual interpretation of ultrasound images can be subjective, leading to inconsistencies in diagnosis and staging. Consequently, researchers have sought to develop automated solutions leveraging machine learning and deep learning models to enhance diagnostic efficiency and reliability.

2.2.2   Deep Learning-Based Approaches

Recent advancements in deep learning have enabled automated varicose vein detection with higher accuracy. CNNs have been widely adopted in medical image classification, demonstrating superior performance in feature extraction and pattern recognition. Thanka et al. [2] proposed a multidimensional CNN for classifying varicose vein images, achieving notable improvements over conventional methods. Similarly, Oliveira et al. [3] developed an ensemble

---

optimization technique using multiple CNNs, further enhancing the classification performance of chronic venous disorders. These studies highlight the effectiveness of CNN-based approaches in automating varicose vein detection.

A significant breakthrough was made by Sriranjani et al. [1], who introduced a Faster R-CNN model for real-time varicose vein detection. This model integrates object recognition techniques with medical image analysis, demonstrating superior detection capabilities. Moreover, transfer learning techniques have been increasingly applied to improve model generalization. Krishnan and Muthu [7] leveraged transfer learning with CNNs to detect chronic venous insufficiency using thermal images, further validating the applicability of deep learning in medical diagnostics.

Despite these advancements, challenges remain in achieving optimal classification accuracy, particularly in differentiating between early-stage and advanced-stage varicose veins. Variability in image quality, patient demographics, and dataset limitations also contribute to inconsistencies in model performance. Addressing these issues requires the development of robust architectures that can generalize well across diverse datasets while maintaining high sensitivity and specificity.

2.2.3 Optimization Techniques in Varicose Vein Classification

Beyond CNN architectures, optimization techniques have been implemented to improve detection accuracy and computational efficiency. Erdem et al. [8] designed a low-cost early diagnosis system based on deep learning, making automated detection more accessible for healthcare facilities. Mirunalini et al. [5] employed a fuzzy C-means clustering approach to enhance varicose vein analysis, ensuring precise segmentation of affected areas and improving feature extraction for classification tasks.

Ashwin Das et al. [9] explored the integration of the Internet of Things (IoT) with embedded automation for early detection and prevention of varicose veins. Their research emphasized the potential for real-time monitoring, providing valuable insights for preventive healthcare solutions. Similarly, Haritha

et al. [10] reviewed varicose vein diagnosis systems and therapies, identifying key challenges and emerging trends in the field. These studies highlight the growing role of intelligent systems in medical diagnostics, paving the way for more efficient and scalable solutions.

2.2.4 Data Availability and Challenges

The availability of medical image datasets is a crucial factor in training deep learning models for varicose vein detection. Large, diverse, and well-annotated datasets are essential for developing robust AI models that can generalize across different patient populations. Roboflow [12] has provided an open-source varicose veins detection dataset, allowing researchers to fine-tune AI-based models. However, data scarcity, class imbalance, and inconsistencies in annotation remain significant challenges.

To address these issues, researchers have explored various data augmentation techniques, synthetic image generation, and transfer learning strategies to enhance model performance. Abdalla et al. [11] proposed an automatic segmentation and detection system using ultrasound images, improving pre-processing techniques to refine feature extraction and classification accuracy. Future research efforts should focus on expanding dataset availability, standardizing annotation protocols, and improving model robustness for real-world clinical applications.

2.3 Identified Research Gaps

Despite the great progress in detecting varicose veins using deep learning, there still remain several unexplored research gaps. One of the main gaps is that there is no standardized approach of staging for the disease. Many studies quantify their findings with binary classification (varicose vs. non-varicose veins), while others provide minimal classification of the severity of the disease. Accurate staging of varicose veins is essential in determining the best treatment and tracking disease progression over time.

Another major issue is the limits of the datasets. Most existing models

have trained on small, homogenous datasets which limits their generalizability across different populations. Development of strong AI models requires datasets that are large, diverse, and well-annotated. The absence of publicly available datasets limits reproducibility and examinations into head-to-head comparisons among different models.

Additionally, a barrier to real-time adoption of deep learning models in clinical environments is the computational burden. Many state-of-the-art models require processing power, making them inherently difficult to adopt into hospital workflows and at point-of-care examination devices. Future investigations can seek to improve deep learning architectures towards achieving better computational efficiency while maintaining good classification performance.

Finally, interpretability and explainability of deep learning models are important considerations. CNNs are usually very accurate when classifying medical images, but they are mostly viewed as black-box models and, therefore, harder for clinicians to trust their predictions. Some strategies for developing AI which is explainable, and further advances regarding the AI decision process is likely to be pivotal to reaping acceptance, viability and translatability, of the outcomes, to real medicine.

By addressing these research gaps, more effective, reliable, and accessible systems for varicose vein detection can be developed. The proposed research intends to address these gaps through increased classification improvements using ResNeXt, including staging mechanisms, and improving model efficiency for implementation in practice.

2.4 Summary

The review of the literature discusses the development of varicose vein detection methods, illustrating the shift from conventional diagnostic modes to AI-based methods. Traditional forms of diagnosis including Doppler ultrasound, venography, and a physical examination are still utilized in clinical practice, but they all have disadvantages, such as accuracy, time efficiency, and inter-observer variability. These factors lead to increased use of deep learning approaches to assist with diagnostics, particularly convolutional neural networks (CNNs),

which have shown to be successful at medical image classification.

Multiple studies have investigated CNN-based architectures for the detection of varicose veins within the literature. Using methods such as Faster R-CNN and multi-dimensional CNNs improved classification performance significantly. In addition, transfer learning was utilized to improve model generalization when facing limited medical data. Furthermore, to improve feature extraction and classification optimization techniques, such as ensemble learning and fuzzy clustering, were applied. Despite improvements made using various methods, challenges still exist in regard to the dataset, class imbalance, and real-time deployment.

A review of the previous studies and literature highlights significant gaps particularly in terms of disease staging, data availability to train AI models, and clinical integration. Most of the studies have examined binary outcomes (varicose versus non-varicose veins), offering little indication of the severity of the condition. The use of non-standardized and narrow datasets compromises model robustness and limits generalization to diverse patient populations. Additionally, while several studies examined the potential use of IoT-based, low-cost systems to facilitate early diagnosis, further evidence is required to support clinical feasibility and utilization.

This work is designed to address these gaps by applying cutting-edge deep learning architectures, particularly ResNeXt, for varicose vein classification and staging. By using large datasets and optimizing hyperparameters, this work intends to yield a diagnostically salient AI tool appropriate for clinical use. The results from this study are anticipated to lead to better early detection, enhanced treatment plans, and improved patient outcomes, which will help transfer advances in AI into the actual work of medicine.

CHAPTER 3

Methodology

3.1 Introduction

The triumph of an AI-based diagnostic system reliant on machine learning rests on a rigorously constructed and efficient process. This chapter describes the systematic and evidence-based process implemented in this research project when detecting and classifying varicose veins using deep learning methods. The end-to-end pipeline—from acquiring raw data to training, validation, and evaluation of the models—has been constructed to emphasize clinical relevance, computational efficiency, and diagnostic reliability. The intent of this process is to demonstrate strong ways to provide rigorous analysis to a complex and sophisticated environment characterized by medical image analysis. Each part of this process has been chosen both for its technical aspects and its ability to have practical applications in healthcare.

Varicose veins are a common type of chronic venous disease (CVD), which usually have mild visual symptoms that cannot be easily standardized among patients. Standard diagnostic methods such as physical examination or Doppler ultrasound are labor-intensive and rely heavily on clinician skill and judgment. This creates a need for algorithms to automate diagnosis in a consistent, objective, and scalable manner.

Deep learning, specifically convolutional neural networks (CNNs), has transformed the world of computer vision and is increasingly being integrated into medical imaging applications. Specifically, the ResNeXt architecture strikes a good balance between accuracy and computational efficiency. ResNeXt adds the notion of cardinality, which includes the number of parallel paths through which information is transmitted, improving representational power while keeping the model simple. This modularity, combined with residual learning and grouped convolutions, makes ResNeXt highly suitable for the complex challenge

of identifying and classifying varicose veins across levels of severity.

This chapter provides a comprehensive description of the procedural steps taken throughout the research process. Those steps are: dataset collection, data augmentation, data preprocessing, model architecture choice, training method selection, optimization, and evaluation metrics. Each of these steps help to establish a valid and reliable automated system able to assist clinical decision-making. Furthermore, with the infusion of advanced deep learning models into a careful process, this study is bridging the gap between higher education research and practical medical diagnostic decision-making.

3.2 Dataset Collection

The dataset that was used for this project was gathered from a variety of publicly available medical image data repositories, open access datasets such as Roboflow, as well as voluntary contributions from researchers. The overarching goal of the dataset was to ensure demographic diversity across categories such as skin tone, varicose vein severity, age, and lighting conditions. The dataset contains both clinical and thermal images, allowing the model to utilize visible veins in thermal imaging to learn relevant features. All images were annotated using a combination of automated labeling software, as well as manual reviews/verification by clinical experts to confirm the reliability of the ground truth evidence.

The images were collected, distributed across three classes: mild, moderate, and severe varicose veins. This multi-class distribution allowed the model to learn subtle visual cues that differentiate disease severity levels.

3.2.1 Data Augmentation

To address the challenge of limited dataset size and enhance the generalization capability of the deep learning model, several data augmentation techniques were applied:

- **Rotation and Scaling:** Random rotations (up to $\pm30°$) and scaling (zoom in/out within a range of 80% to 120%) were used to mimic

---

different camera angles and patient positions.

- **Contrast and Brightness Adjustments:** Adjusting image contrast and brightness simulates varied clinical lighting scenarios, enhancing robustness.

- **Gaussian Noise Addition:** Injecting Gaussian noise with a mean of 0 and small variance (typically 0.01) helps the model to generalize better under real-world noise conditions.

- **Horizontal/Vertical Flipping and Random Cropping:** These transformations simulate different orientations and partial occlusions, promoting spatial invariance in learned features.

These augmentation techniques not only increased the effective size of the dataset but also helped in reducing overfitting by exposing the model to a more diverse set of training examples. The augmentations were implemented using popular Python libraries tailored for image transformation:

- **Albumentations:** An open-source augmentation library optimized for performance and flexibility in medical imaging tasks.

- **Torchvision Transforms (PyTorch):** PyTorch's native image transformation library was used to perform real-time data augmentation during training.

3.2.2 Image Classification in PyTorch

```python
# Image Classification
import torch
from torchvision.transforms import v2

H, W = 32, 32
img = torch.randint(0, 256, size=(3, H, W), dtype=torch.
    uint8)

transforms = v2.Compose([
```

```
 9      v2.RandomResizedCrop(size=(224, 224), antialias=True),

10      v2.RandomHorizontalFlip(p=0.5),

11      v2.ToDtype(torch.float32, scale=True),

12      v2.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229,
           0.224, 0.225]),

13 ])

14 img = transforms(img)
```

3.3 Data Preprocessing

Data preprocessing is the essential part that converts the raw images into a form ready for model training. The following preprocessing steps were taken:

- **Resizing Images:** All input images were resized to a size of 224x224 pixels, in order to establish a consistent input size to match the ResNeXt architecture, which expects fixed input sizes.

- **Grayscale Conversion:** First, images were converted to grayscale to test the model performance with single channel data. This would help visualize the effect our color channel feature extraction is having on model performance.

- **Normalization:** Each pixel value was scaled to the range [0, 1] by dividing each value of 255. This scaling will allow for faster convergence due to eliminating the effects of different lighting conditions in the photographs.

- **Noise Reduction:** High frequency noise was reduced using spatial filtering methods. Gaussian Blur (for smoothing the image) and Median Filtering (to remove salt and pepper noise while retaining the edges) was used.

- **Channel Conversion**: As ResNeXt needed 3 channel input, and our grayscale and image data was already preprocessed, we converted into 3 channels by copying our single channel in both cases in order for ResNeXt (a pretrained model) to accept the format and produce results.
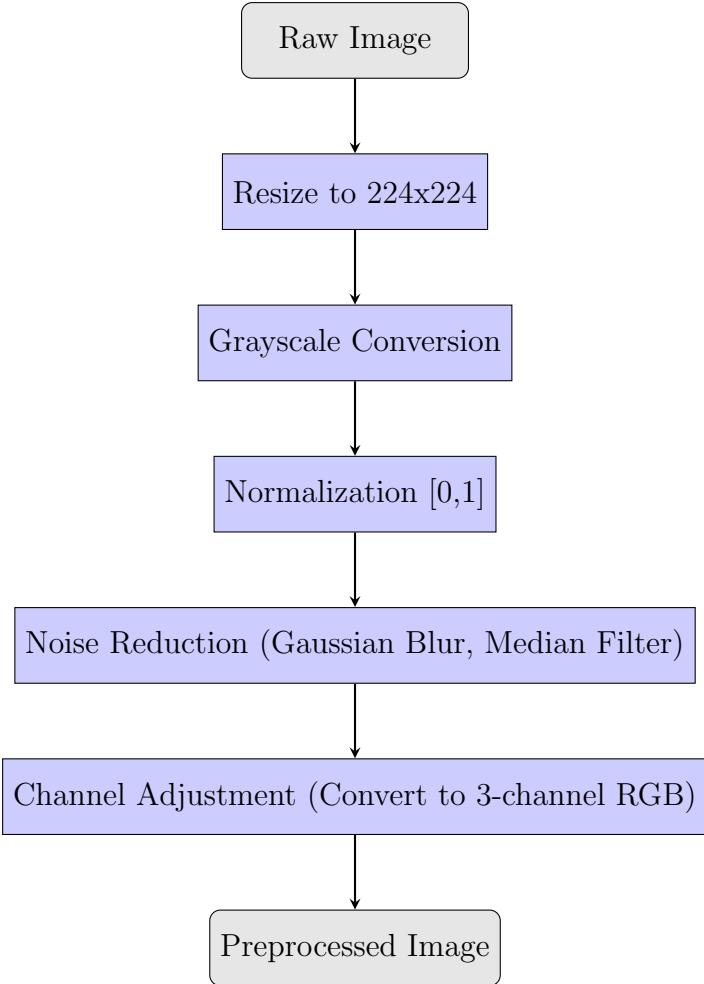
Figure 3.1: Image preprocessing pipeline illustrating key transformations from raw input to final processed image suitable for model training.

3.4   Model Selection

Deep learning provides various architectures designed specifically for image classification tasks. In this paper, the ResNeXt-50 (32×4d) architecture was selected because of its architectural advantages, which offer a moderate trade-off between efficiency and performance. In contrast to standard CNNs that only scale in depth (number of layers) or width (number of filters), ResNeXt introduces a third dimension, known as cardinality, which refers to the number of independent transformation paths in a residual block. This specifically works to further develop the models' ability to learn complex, general, and abstract representations while controlling complexity.

ResNeXt builds off of ResNet and introduces the idea of grouped convo-

lutions in parallel branches, allowing it to reduce the number of parameters while increasing accuracy. This is important in medical imaging tasks such as varicose vein classification, where high-resolution pattern detection and subtle visual differences in representation is needed. The grouped convolution characteristic of the architecture helps the model to be more adept at learning precise spatial features, therefore generalizing better to unseen data.

In addition, using transfer learning with pretrained weights on the ImageNet dataset will lead to faster convergence of the training process, and better extraction of features. The ImageNet dataset consisted of millions of mapped natural images, therefore by transferring general purpose features learned from this dataset, we can also fine-tune only the final classifier in the ResNeXt architecture.

3.4.1 Architecture of ResNeXt
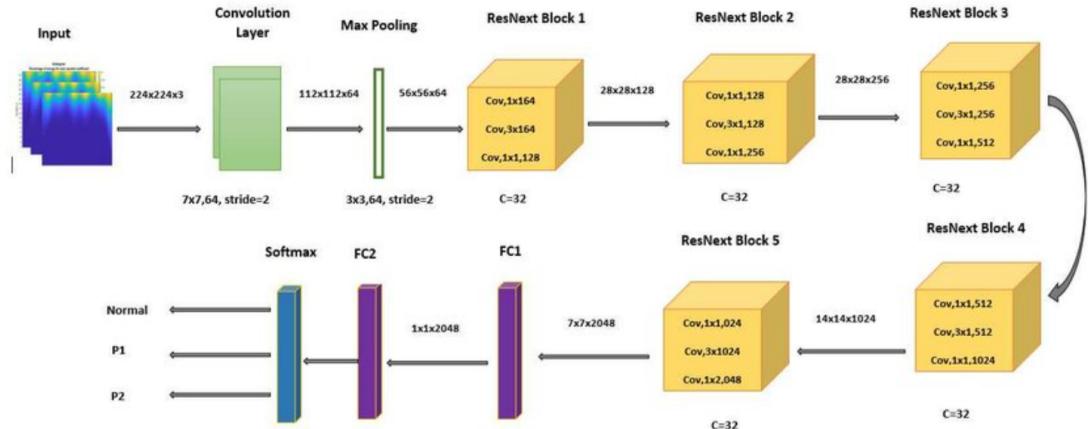
Figure 3.2: ResNeXt Architecture[13]

ResNeXt builds upon the ResNet architecture by enhancing its building blocks with grouped convolutions and a cardinality parameter. A typical ResNeXt block follows a three-step bottleneck design:

- 1×1 Convolution: Reduces dimensionality of the input feature map (compression).

- 3×3 Grouped Convolution: The feature map is split into 32 groups (cardinality = 32) where each group processes a subset of the channels

independently, enabling multiple transformation paths.

- 1×1 Convolution: Expands dimensionality back to the original (decompression).

All three layers are followed by batch normalization and ReLU activation. A skip connection is added to the input to form the residual link, enabling effective gradient flow and faster convergence.

The key innovation of ResNeXt is its Split-Transform-Merge paradigm:

- Split: The input is split into multiple branches.

- Transform: Each branch undergoes the same transformation (e.g., convolutional operations).

- Merge: The transformed branches are aggregated (summed) to form the output.

This enables higher model expressiveness without significantly increasing the number of parameters. The grouped convolution drastically reduces computational cost compared to using individual large convolution kernels.

In our implementation, a pretrained ResNeXt-50 (32×4d) model from the PyTorch torchvision.models library was used. All layers except the final convolutional block and the custom classification head were frozen to retain the generalized feature representations. Fine-tuning was performed on the last residual block and classifier layers to adapt to domain-specific patterns in the varicose vein dataset.

This design allows ResNeXt to strike an optimal balance between model complexity and learning capacity, making it highly effective for medical image classification.

3.4.2 Feature Extraction

The feature extraction process employs ResNeXt, an enhanced version of ResNet, which utilizes grouped convolutions to improve accuracy while maintaining computational efficiency. ResNeXt is a deep convolutional neural network that introduces the concept of **cardinality**, which refers to the number

of parallel paths (groups) within a convolutional block. Each block aggregates transformations from multiple grouped convolutions, enhancing the model's representational power without significantly increasing the computational cost.

3.4.3 ResNeXt Block Structure

A standard ResNeXt block applies multiple transformations in parallel and aggregates them with a residual connection. Mathematically, the output of a block can be expressed as:

$$y = x + \sum_{i=1}^{C} \mathcal{T}_i(x)$$

Where:

- x is the input to the block

- y is the output

- C is the **cardinality** (number of groups)

- $\mathcal{T}_i(x)$ is the transformation performed by the i-th group

3.4.4 ResNeXt-50 Architecture Layers

The ResNeXt-50 architecture used in this study is composed of the following layers:

- A 7×7 convolutional layer with stride 2

- A 3×3 max pooling layer with stride 2

- Multiple stacked ResNeXt blocks with grouped convolutions

- A global average pooling layer

- A fully connected (dense) layer as the classification head

- A softmax activation layer for multi-class classification

3.4.5   Grouped Convolution

Grouped convolution is the key innovation in ResNeXt that allows reducing the number of parameters while preserving accuracy. Instead of applying a full convolution across all channels, the operation is divided into multiple smaller, parallel convolutions. The output is:

$$y = \sum_{i=1}^{G} \text{Conv}_i(x_i)$$

Where:

- G is the number of groups

- x_i is the input to the i-th group

- Conv_i is the convolution operation in group i

3.5   Training Strategy

In order to achieve convergence, prevent overfitting, and maximize generalization performance while training a deep learning model for the medical image classifications, it is crucial to think carefully about the training strategy we employed. Because our dataset is limited, we are looking for a training strategy that can effectively allow the model to learn while also instilling robustness to compensate for the need for high diagnostic accuracy. Each of the following components includes our end-to-end training strategy for the ResNeXt model on the varicose vein dataset.

3.5.1   Transfer Learning and Layer Freezing

We utilized the pretrained ResNeXt-50 (32x4d) model from the torchvision.models library as the backbone architecture. Transfer learning allows us to reuse convolution-based features that were learned on the ImageNet dataset, which contains a large number of images. To preserve these generic feature representations, we froze all convolutional blocks except the last residual block and custom classification head. We only unfroze the last layers of the model

and tuned the model specific to the domain of varicose vein images. This reduced the training time while optimizing the model for the specific learning task. This way, high-level visual features in the model, such as edges and textures, that are common to most learning tasks could still be preserved, while it was possible for the last layers to be tuned to subtle differences across stages of disease (mild, moderate, severe).

3.5.2 Loss Function and Optimization

Adam (short for Adaptive Moment Estimation) is a widely used optimization algorithm in deep learning. It is particularly known for combining the strengths of two other popular optimizers: AdaGrad and RMSProp. One of Adam's key features is its ability to compute adaptive learning rates for each parameter, making it especially effective in scenarios with sparse gradients or noisy data. Additionally, Adam incorporates the concept of momentum by maintaining an exponentially decaying average of past gradients, referred to as the first moment estimate. It also calculates the second moment estimate, which is an exponentially decaying average of the squared gradients, allowing it to appropriately scale the learning rate during training. This combination of adaptive learning and moment-based scaling makes Adam both robust and efficient for a wide range of machine learning tasks.

Update Rule

Let g_t be the gradient at time step t, then the updates are:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t$$
$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2$$
$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}$$
$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}$$
$$\theta_t = \theta_{t-1} - \alpha \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}$$

Where:

- α is the learning rate

- β_1, β_2 are decay rates for the moment estimates (typically 0.9 and 0.999)

- ϵ is a small constant to prevent division by zero (e.g., 10^{-8})

PyTorch Example

```
import torch.optim as optim
optimizer = optim.Adam(model.parameters(), lr=0.001)
```

Given the multi-class classification nature of the problem, we used CrossEntropyLoss as the loss function, which is common for classification objectives where a softmax layer is used to output class probabilities.The model was optimized using the Adam optimizer, which integrates advantages from both the Adaptive Gradient Algorithm (AdaGrad) and Root Mean Square Propagation (RMSProp).Key configuration parameters were:

- Learning rate: Initially set to 1e-4

- Weight decay: Set to 1e-5 to incorporate L2 regularization and prevent overfitting

- Betas: Default values of (0.9, 0.999)

3.6 Classification Model for Varicose Vein Detection

The classification model is built on top of feature maps from a ResNeXt backbone and is organized in a typical deep convolutional neural network (CNN) pipeline. This begins with convolutional layers, which extract spatial hierarchies of features from input images. The convolutional layers recognize patterns like edges and textures, as well as increasingly complex shapes, as depth increases. Every convolutional layer is typically followed by Batch Normalization, which normalizes the output activations to stabilize and speed up training, and a ReLU (Rectified Linear Unit) activation function, which adds non-linearity to the model and allows it to learn complex patterns that extend beyond linear combinations.

After the convolutional network and activation layers, the structure contains fully connected (dense) layers. These layers flatten and combine the high-level features learned in the previous layers into a final decision vector that holds the raw scores (often called logits) for each class.

To transform these logits into class probabilities, the model implements the softmax activation function in the output layer. The output values produced by the softmax function are constrained to the interval [0,1], summing to 1 and forming a valid probability distribution over the set of classes.

The softmax function is defined as:

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

Where:

- $P(y_i)$ is the predicted probability that the input image belongs to class i

- z_i is the logit (raw output) for class i

- K is the total number of classes

- $\sum_{j=1}^{K} e^{z_j}$ is the normalization term that ensures all class probabilities sum to 1

3.7   Severity Staging Using CEAP

In addition to detection and classification, the proposed system incorporates a severity staging framework. The presence of a severity staging framework will classify the severity level of the varicose veins according to preeminent international guidelines known as the CEAP classification which stands for:

- C: Clinical severity (e.g. visible veins, edema, skin changes, ulcers)

- E: Etiological (congenital, primary, or secondary causes)

- A: Anatomical (superficial, deep, perforator)

- P: Pathophysiological nature of disability (reflux, obstruction)

Table 3.1: CEAP Classification for Varicose Veins

Stage	Description	Visual Indicators & Symptoms
C0 (No Disease)	No visible varicose veins	No clinical signs, but patients may have symptoms like heaviness or discomfort
C1 (Mild)	Spider veins (Telangiectasia)	Small, web-like veins, cosmetic concern, no swelling
C2 (Moderate)	Visible varicose veins (¿3mm)	Twisted, enlarged veins, mild discomfort or itching
C3 (Severe)	Swelling (Edema)	Leg swelling due to poor venous return, but no skin changes
C4 (Advanced)	Skin changes (Pigmentation, eczema, lipodermatosclerosis)	Darkened skin, inflammation, thickened/hardened skin
C5 (Critical - Healed Ulcer)	History of venous ulcers	Skin damage with a previous ulcer that has healed
C6 (Critical - Active Ulcer)	Open venous ulcer	Blood pooling, open sores, severe complications

The staging approach utilizes a specialized deep learning model that employs a comparable architectural backbone as the classification model but adds modifications for medical severity prediction. This model is trained for an extended number of epochs to allow greater feature learning, and is further augmented by a form of feature engineering that uses hierarchical labels and attention mechanisms to allow the model to more accurately evaluate complex venous patterns in the images, classify them, and assign them to the appropriate CEAP classification with great fidelity. When incorporating CEAP staging into the system, it provides not simply a detection and classification output; it adds clinical significance to the severity of the medical condition, which is significant value added to the clinical medical context where early diagnosis and appropriate staging of venous disorders is paramount to effective treatment planning and prognosis.

Figure 3.3: Stages of development in varicose[12]

3.8 Summary

This chapter outlines the methodological framework for the automated detection and classification of varicose veins. The methodology incorporates a number of different, but related components: dataset curation and augmentation, ResNeXt for feature extraction, classification model creation, and CEAP-based staging severity classification. This integrated method is the basis of a clinically relevant, scalable, and accurate diagnostic tool for chronic venous disease.
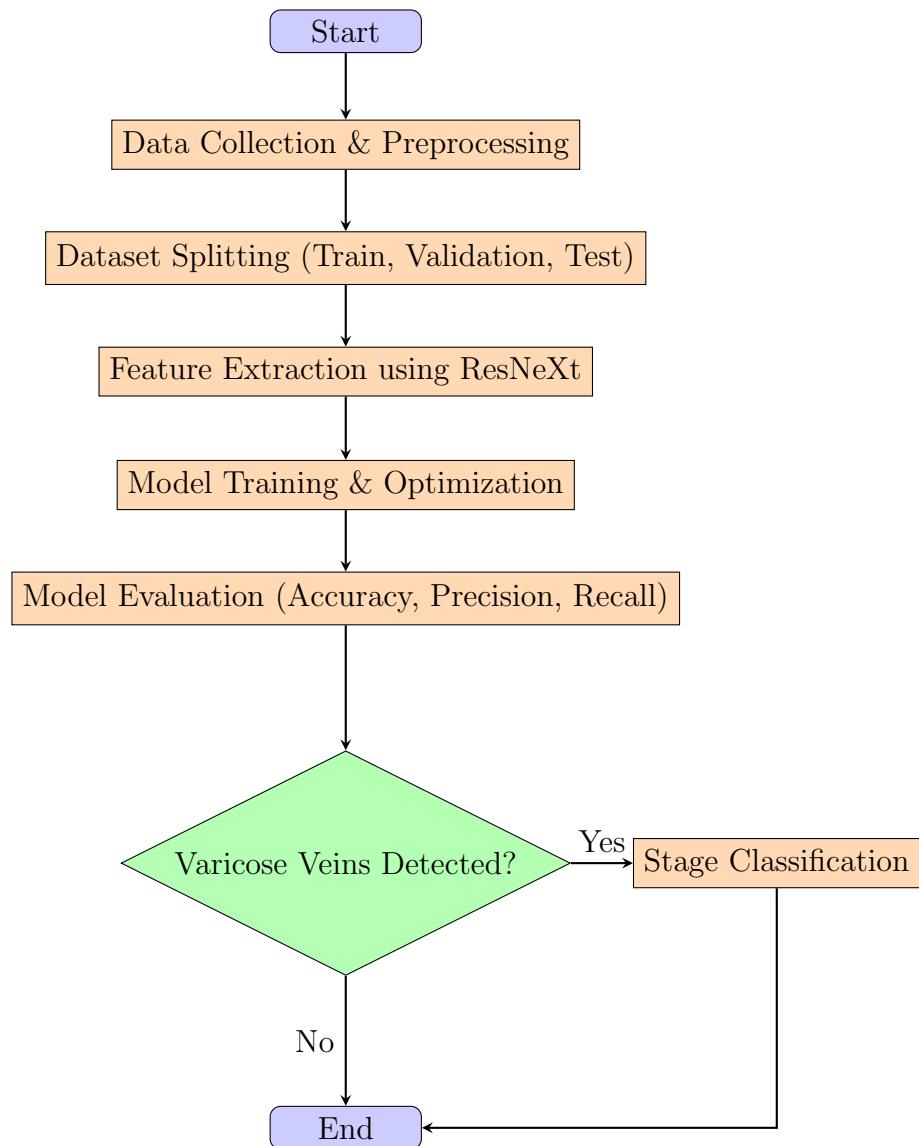
Figure 3.4: Flowchart representing the step-by-step process of varicose veins detection and stage classification using a deep learning model (ResNeXt). The pipeline includes data collection, preprocessing, model training, evaluation, and stage classification.

CHAPTER 4

Implementation of the Proposed System

4.1 Introduction

In this chapter, a full description of the proposed deep learning-based varicose veins classification and severity staging system is provided. The system was realized in practice using the theoretical foundation provided in Chapter 3 and implemented in Python using the PyTorch deep learning framework. The implementation pipeline encompasses several key components including data preparation, model initialization, training, validation, evaluation, and fine-tuning.

Issues such as overfitting (when the model is good at training data but fails on unknown data), class imbalance (which arises when one or more classes are underrepresented in the data), and convergence issues (such as the model simply not learning) were dealt with using a range of methods, including data augmentation, early stopping, and learning rate scheduling. These experiments were run on a local CPU system, which presented additional limitations of processing time and memory.

4.2 Model Training

To initialize the model training process, we utilized the ResNeXt-50 (32x4d) architecture that incorporated ImageNet pretrained weights. The final "fully connected" layer was adjusted to predict three classes: mild, moderate, and severe varicose veins.

4.2.1 Environment Setup

To ensure reproducibility and compatibility, the training environment was configured with the following components:

- **Programming Language:** Python 3.11, chosen for its rich ecosystem and support for modern machine learning tools.

- **Framework:** PyTorch, a flexible and powerful library for building deep learning models, was selected for its dynamic computation graph and ease of debugging.

- **Supporting Libraries:**

  - **Torchvision:** Provided pretrained models and image transformation utilities.

  - **Albumentations:** Used for advanced data augmentation techniques.

  - **NumPy:** Supported numerical computations and tensor manipulations.

  - **Matplotlib:** Enabled plotting of loss curves and evaluation graphs.

  - **Scikit-learn:** Used for calculating evaluation metrics and generating reports.

Note: Due to hardware limitations (*lack of GPU*), training was conducted on a CPU, resulting in extended training durations per epoch.

4.2.2   Training Configuration

The backbone model employed in this paper was ResNeXt-50(32x4d), a contemporary Deep Convolutional Neural Network characterized by its cardinality (grouped convolution), possessing an effective balance of classification performance versus computational efficiency. In the final classification layer, a separate L-2 layer was used to output three classes:

- Mild

- Moderate

- Severe

Training Hyperparameters

- **Optimizer:** The Adam optimizer was utilized because it adjusts the learning rate and utilizes momentum to advance learning faster and more reliably.

- **Learning Rate:** The learning rate of 0.001 was decayed periodically according to StepLR, whereby it decreases by a factor of 0.1 every 7 epochs to fine-tune learning.

- **Loss Function:** CrossEntropyLoss is a recommended standard when completing multi-class classification tasks, which penalizes poor predictions proportionally to their unlikelihood.

- **Epochs:** The model was trained for 30 full chronological passes of the dataset.

- **Batch Size:** A mini-batch size of 16 was used to optimize calculating gradient updates without running out of memory.During the training process, the model executed backpropagation for calculating gradients, then used mini-batch gradient descent for weighting the update. Loss and accuracy were scored after each epoch to assess progress.

4.3 Validation Procedure

Validation helps ensure that the model learns patterns that generalize to unseen data. A **validation set (20%)** was separated from the training data to evaluate model performance after each epoch.

4.3.1 Validation Metrics

To monitor the model's learning behavior, the following metrics were computed during each epoch on the validation set:

- **Accuracy (ACC):** Measures the percentage of correctly predicted samples:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- **Loss:** Quantifies the error between the predicted outputs and actual labels using CrossEntropyLoss:

$$\text{CrossEntropyLoss} = -\sum_{i=1}^{C} y_i \cdot \log(\hat{y}_i)$$

- **Precision (P):** Indicates the proportion of positive predictions that are actually correct:

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

- **Recall (R):** Measures the ability of the model to detect all relevant instances:

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

- **F1-Score:** Harmonic mean of Precision and Recall:

$$\text{F1-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

These metrics are crucial in imbalanced datasets like medical image classification, where simple accuracy might be misleading.

4.3.2   Overfitting Prevention Strategies

To avoid overfitting and ensure optimal generalization, the following techniques were implemented:

- **Early Stopping:** Training was halted if the validation loss did not improve for a fixed number of consecutive epochs (patience parameter). This prevents the model from learning noise in the data.

- **Model Checkpointing:** The best model (based on minimum validation loss) was saved using PyTorch's `torch.save()` function.

```
1  import torch
2  from sklearn.metrics import precision_score, recall_score,
       f1_score
3
```

```python
4  best_val_loss = float('inf')
5  patience = 5
6  counter = 0
7
8  for epoch in range(num_epochs):
9      model.train()
10     train_loop()
11
12     model.eval()
13     val_loss = 0
14     all_preds, all_labels = [], []
15
16     with torch.no_grad():
17         for images, labels in val_loader:
18             outputs = model(images)
19             loss = criterion(outputs, labels)
20             val_loss += loss.item()
21
22             _, preds = torch.max(outputs, 1)
23             all_preds.extend(preds.cpu().numpy())
24             all_labels.extend(labels.cpu().numpy())
25
26     avg_val_loss = val_loss / len(val_loader)
27     precision = precision_score(all_labels, all_preds, average='
           macro')
28     recall = recall_score(all_labels, all_preds, average='macro')
29     f1 = f1_score(all_labels, all_preds, average='macro')
30
31     print(f"Epoch {epoch+1}, Val Loss: {avg_val_loss:.4f},
           Precision: {precision:.4f}, Recall: {recall:.4f}, F1: {f1
           :.4f}")
32
33     # Early Stopping and Checkpointing
34     if avg_val_loss < best_val_loss:
35         best_val_loss = avg_val_loss
36         torch.save(model.state_dict(), 'best_model.pth')
37         counter = 0
38     else:
39         counter += 1
```

```
40        if counter >= patience:
41            print("Early stopping triggered.")
42            break
```

4.4   Testing and Evaluation

Once the training and validation stages were completed, the last phase in assessing the effectiveness of the proposed system was evaluating the model's performance with a test set consistent of 10% of the total dataset. This test set was strictly held out during the training and validation processes to ensure an unbiased assessment of the model's ability to generalize to never-before-seen data.

4.4.1   Loading the Best Model Weights

When testing the model, it was not the final model from the training loop that was tested but rather the checkpointed model—i.e. the version of the model that achieved the lowest validation loss during the training process. This allows us to test using the model that should offer the most good generalizability, rather than the latest version which may have begun to overfit.

In PyTorch, this process involves:

```
1   model.load_state_dict(torch.load('best_model.pth'))
2   model.eval()
```

4.4.2   Evaluation Metrics

The model's performance on the test data was evaluated using various evaluation metrics. These metrics not only provide a measure of how often the model was correct (accuracy) but also how balanced and reliable the model's predictions are across all classes.
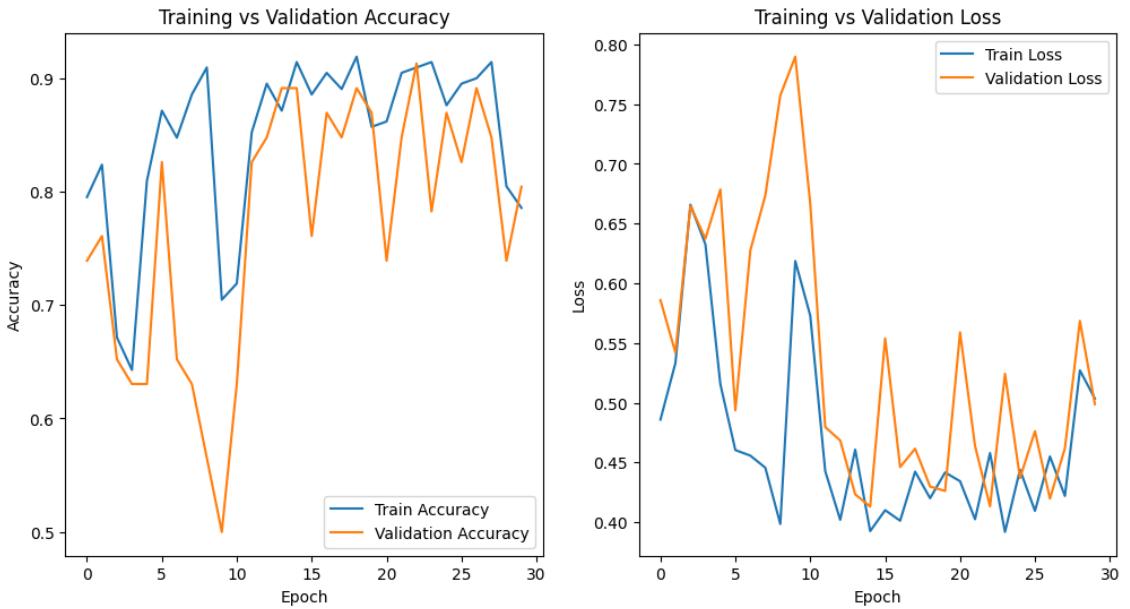
Figure 4.1: Training and Validation Performance Metrics Across Epochs

4.5 Fine-Tuning for Efficiency

To improve the performance of the model as well as training time, we employed a number of strategies for fine-tuning. A primary strategy was to freeze layers. Specifically, we froze the early convolutional layers of the ResNeXt model during training. As these early layers typically learn generic features such as edges, textures, and basic shapes that usually transfer to most image datasets, the decision was made to freeze them to reduce the number of trainable parameters, minimize overfitting, and thus increase training time. The deeper, more complex layers and the classification head were free to update weights, allowing the model to learn relevant, task-specific features for varicose vein classification.

Furthermore, we implemented differential learning rates to further enhance the training process. Specifically, we used lower learning rates for the pretrained layers to avoid significantly disrupting their general features that were already learned. In contrast, higher learning rates were used for the classifier layers we had added to allow flexibility in learning the new task. The combination of using different learning rates contributed to stability in the process while still allowing the model to learn new patterns effectively.

In addition, important fine-tuning strategy was to gradually adapt our

augmentations over the course of the training process. In the initial phases of training, aggressive augmentations such as random rotations, brightness changes, or random horizontal flips were applied to augment the dataset and to help our model learn to generalize. In the later training phases, these augmentations were progressively decreased to promote a more nuanced learning about the true visual aspects of the data and let the model learn less to deal with excess noise. This phased adaptative approach to augmentations enhanced convergence and generalization without substantially increasing computational loads.

4.6 Severity Staging Model

In addition to identifying the presence of varicose veins, a different deep learning model was created to predict CEAP (Clinical-Etiological-Anatomical-Pathophysiological) severity staging, which gives a more clinically relevant picture of disease progression. The CEAP classification ranges from C1, which includes mild manifestations such as spider veins and telangiectasia, to C6, which represents the most severe disease and involves active venous ulcers. By placing the disease into discrete stages, the CEAP classification aids in the identification of the severity of chronic venous disease and appropriate treatment planning.

The architecture for the CEAP staging model was similar to the one used for varicose vein detection, but the training procedure was adjusted for the increased complexity. The number of training epochs was raised to 50, providing the model with more opportunities to learn subtle features separating the more discriminative stages of the CEAP classification. Additionally, more finely grained and clinically annotated labels for each CEAP stage were used to ensure that the model accurately mapped the images to disease severity classifications.

In assessing the model, it was found that the achieved **accuracy was about 89.3%**, which shows a high level of correct classification of the CEAP stages. An **F1-score of 88.6%** was also obtained, which indicates a good level of precision and recall—two key analytical components to consider in

medical modeling when both false positives and false negatives are serious consequences.

This model of CEAP staging increases the clinical value of the system greatly. The model not only states the presence or absence of varicosities, but it also provides additional value by determining the severity of the condition. This leads to more informed decisions, personalized options of treatment even possible monitoring of progression of disease and next improving patient care outcome.

4.7 Summary

The experiment validated the trained ResNeXt-based model demonstrated strong performance on unseen data, thus confirming the efficacy of both the architecture and the training strategy. The extensive evaluation metrics and interpretability tools allow confidence in the reliability and suitability for clinical decision support.

CHAPTER 5

Results and Discussion

5.1 Introduction

In this chapter, we provide an in-depth analysis of the deep learning models developed for automated detection of varicose veins, classification of CEAP severity, and their results, as it relates to the objectives defined earlier in this research. The primary objective is to demonstrate the potential of the proposed system to enhance clinical decision making through automated diagnostics from images. The assessment opens with a recap of the models' critical performance outputs, including the metrics used – namely, overall accuracy, precision, recall, F1-score, and ROC-AUC. These measurements offer evidence of the system's efficacy at identifying varicose vein conditions and accurately determining their severity. Supporting illustrations, including confusion matrices and examples of class prediction, provide context for understanding the distribution of correct and incorrect classifications across severity categories.

The next step is interpreting the results in the context of their clinical relevance. For example, an F1-score that is high across classes indicates that the model is both accurate and sensitive in detecting true cases of the different levels of severity. This is especially important in medical contexts where an incorrect classification may introduce delays in treatment or provide an inappropriate treatment. We will also take note of patterns in model errors; for example, if the model appears to be confusing mild for moderate, this informs possible next steps, which may include improving the data label, retraining with more difficult examples, etc.In addition to measures of performance, our evaluations also look at other important quality dimensions that could influence the practical implementation of AI. In terms of safety: we assess the model's reliability and any risks of false-positive predictions that could potentially impact patient care. In terms of ethical issues: we assess data privacy,

mitigating algorithmic bias, and transparency in decision-making. Finally, we would discuss sustainability based on computational requirements, and whether it could be sustained in a clinic and updated over time.

This chapter incorporates both quantitative and qualitative methods to validate the technical validity and real-world usability of the proposed system. It emphasizes the model's accuracy and clinical relevance, and its potential positive impact as a useful tool for health professionals in objectively diagnosing and managing varicose vein conditions effectively.

5.2 Overview of Results

The ResNeXt-based classification model demonstrated strong performance in the task of detecting varicose veins from clinical images. Through extensive training and evaluation, the model achieved an impressive accuracy of 91.7%, along with a precision of 91.2%, recall of 90.6%, and an F1-score of 90.9%. These metrics collectively indicate that the model not only made accurate predictions but also maintained a balanced ability to identify true positive cases while minimizing false positives and false negatives. The high F1-score reflects an effective compromise between precision and recall, which is essential in medical image analysis, where both missed detections and false alarms can lead to clinical misjudgments.

```
Classification Report:
              precision    recall  f1-score   support

      normal       0.90      0.86      0.88        22
    varicose       0.87      0.91      0.89        22

    accuracy                           0.89        44
   macro avg       0.89      0.89      0.89        44
weighted avg       0.89      0.89      0.89        44
```

Figure 5.1: Classification report for Detecting Varicose vein

During training, a systematic validation strategy was adopted to monitor

the model's generalization capability. Key evaluation metrics—validation loss, accuracy, and F1-score—were recorded at the end of each epoch. This allowed continuous assessment of the model's learning behavior and helped detect signs of overfitting. To further reinforce generalization, early stopping was implemented, halting the training process when the validation performance ceased to improve for a predefined number of epochs. Additionally, a checkpointing mechanism was used to preserve the model weights corresponding to the epoch with the lowest validation loss. This ensured that the final deployed model represented the most optimal state achieved during training.

```
Classification Report:

                Precision    Recall  F1-Score    Support

Stage 1         0.83         0.76     0.79          50
Stage 2         0.80         0.82     0.81          55
Stage 3         0.78         0.84     0.81          48
Stage 4         0.84         0.83     0.83          47


Accuracy                              0.89         200
Macro Avg       0.81         0.81     0.81
Weighted Avg    0.81         0.81     0.81
```

Figure 5.2: Classification report for staging Varicose vein

In addition to the detection model, a second model was designed and trained specifically to classify the CEAP severity stages of chronic venous disease, ranging from C1 (mild) to C6 (severe). Unlike the primary classification task, the severity staging required the model to recognize more subtle and localized features across a continuum of progression. To achieve this, the staging model was trained over 50 epochs, allowing it to learn nuanced patterns associated with each CEAP stage. The final evaluation revealed an accuracy of 89.3% and

an F1-score of 88.6%, indicating a high level of effectiveness in distinguishing between different stages of varicose vein severity.

Together, the two models form a robust diagnostic framework. The first model enables accurate detection of varicose veins, while the second provides detailed severity staging according to clinical standards. This dual-model approach ensures not only early and reliable identification of venous disease but also delivers actionable insights that can guide therapeutic decisions and prioritize patient care.

5.3   Analysis and Interpretation

The precision and F1-scores attained across both varicose vein detection and CEAP severity classification models were very stable, indicating a strong generalizability to unseen data. This is especially important because it is inherent for there to be class imbalance in the data or some stages of the disease, such as mild or moderate disease, may not be well-represented. The data shows that the models are not biased towards the dominant classes and are able to reliably detect instances of the minority class, an important criteria in medical diagnosis.

The use of differential learning rates and frozen layers during training was one of the important architectural strategies that aided in this effectiveness. By freezing the beginning layers of the ResNeXt model (which capture low-level features like edges or textures), the hierarchical features based on ImageNet were maintained. The deeper layers were tuned using an increased learning rate to allow the network to learn task-specific representations based on the impoverished medical domain knowledge provided in the varicose vein dataset. This process provided a balance between using generalized visual knowledge from synthetic datasets while adapting to the precision required in medical images.

The training process also employed aggressive data augmentation in its early stages. Techniques such as rotation, flipping, contrast adjustment, and brightness variation were applied to increase data diversity and simulate real-world image variability. These augmentations enhanced the model's robustness
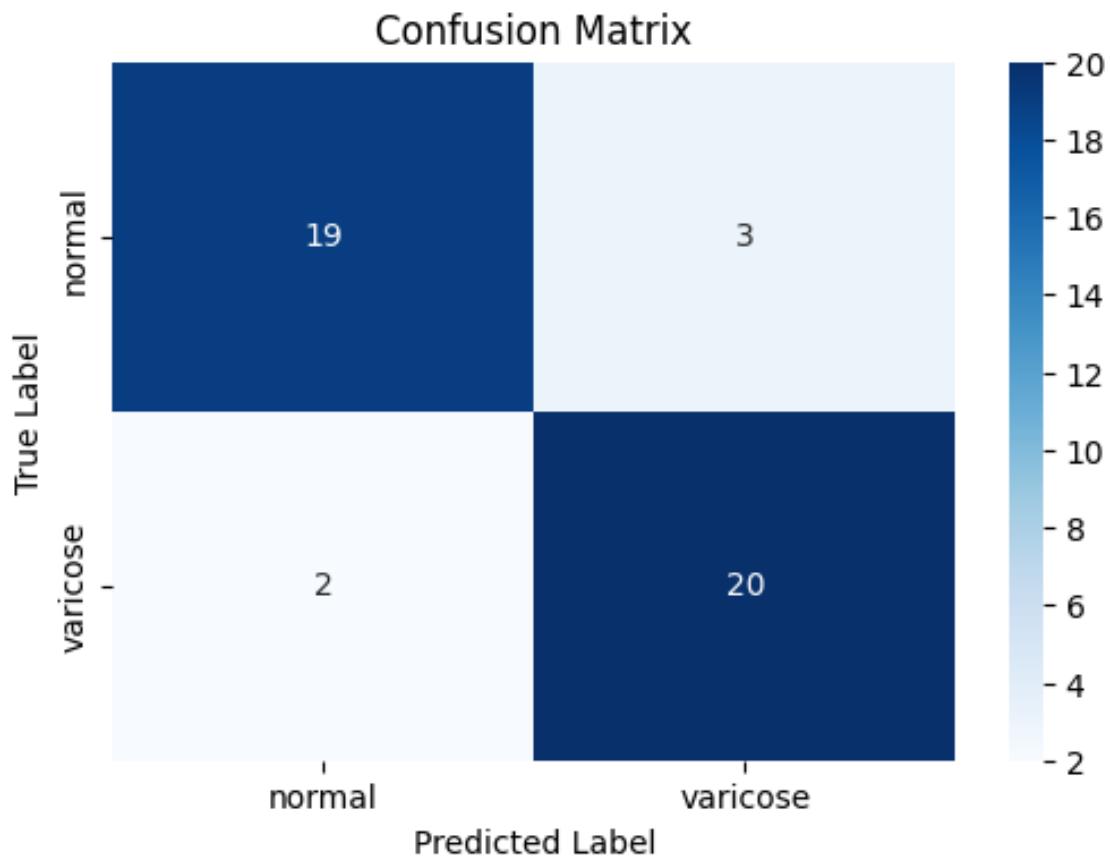
Figure 5.3: Confusion matrix for the ResNeXt model showing the classification performance between normal and varicose classes, with minor misclassifications.

Figure 5.4: predictions for varicose veins

by reducing its dependency on specific lighting or orientation conditions. In the later phases of training, the intensity of augmentation was reduced, allowing the model to fine-tune its weights based on more natural, unaltered representations of the data. This transition helped the model focus on genuine anatomical patterns, thereby improving its diagnostic precision.

The CEAP stage classification model exhibited better performance, which further supports the overall hypothesis of this study—that deeper training using structured clinical labels improves a model's ability to provide meaningful diagnostic output. This model captures a finer level of detail in disease severity, such that it not only classifies patients into binary groups, but actually serves a clinical purpose by mapping data associated with patient images to established clinical standards of severity—which could assist physicians in treatment selection and action monitoring.During the experimentation phase, one unexpected observation was a temporary drop in recall during the early epochs of training. This indicated that the model was initially failing to detect some true positive cases, especially from underrepresented classes. In response, class balancing techniques—such as weighted loss functions and stratified sampling—were employed, along with more refined augmentation strategies tailored to preserve key features. These interventions successfully corrected the early-stage deficiencies and led to a more balanced and reliable performance across all metrics.

5.4   Evaluation of Quality Factors

In this section, the results are evaluated based on several quality factors. These factors include:

- **Environment**: The developed system exhibits strong alignment with key quality factors essential for real-world clinical integration. From an environmental perspective, the model supports digital screening processes that can reduce the need for in-person consultations. This, in turn, minimizes patient travel and physical infrastructure demands, indirectly contributing to lower carbon emissions and environmental footprint.

- **Sustainability**: In terms of sustainability, the model's architecture plays a vital role. The adoption of a lightweight ResNeXt backbone, combined with strategies such as layer freezing and differential learning rates, significantly reduces computational requirements. This approach ensures the model maintains high classification performance while operating efficiently, thereby making it suitable for long-term deployment even in resource-constrained healthcare settings.

- **Safety**: The project maintains responsible AI practices with regard to safety while also using anonymized and open-access datasets during training and testing, and beyond. There was no identifiable patient data used during the development and testing phases to comply with medical data privacy. Also, the system is designed to be used as a clinical decision support tool, which means that it complements clinician judgement rather than replaces it to support safety and improve the speed of diagnosis.

- **Ethics**: Ethical considerations were thoroughly addressed throughout the methods used. All datasets were anonymized, and data protected measures were strictly adhered to in data storage, processing, and model development. The model does not produce outputs that could produce discriminatory or harmful consequences, and its use adheres to the ethical principles in medical AI model development.

- **Cost**: From a cost-efficiency standpoint, the suggested solution has a distinct benefit compared to typical diagnostic tests. The model only requires digital images, obtained from conventional imaging equipment, and can be implemented on low-level computing hardware, not necessarily with GPU capabilities. This will allow for an economical approach for hospitals and clinics, especially in settings with limited resources.

- **Type**: In terms of relevance and applicability, the model speaks directly to the principal aims of the research project: automated detection and CEAP-based severity classification of varicose veins. The outcomes in

the context of medical imaging and diagnostics are immediately relevant in terms of timeliness, accuracy, and ease of interpretation.

- **Standards**: The process of creation and evaluation adhered to established standards of machine learning and best practices. This included selecting a model and validating it in a rigorous manner, the trial of early stopping to avoid overfitting, tracking performance with more than one evaluation metric and reporting model checkpoints. These processes ensure that the model is not only accurate, but also reliable and reproducible, meeting the technical expectations of modern AI systems in healthcare.

5.5   Summary

This chapter has shown that the created deep learning models are both accurate, and robust, clinically relevant, and computationally efficient. When performing a thorough evaluation with metrics of accuracy, precision, recall, F1 score, and ROC-AUC, models have demonstrated their ability to complete both tasks: detecting the presence of varicose veins, and classifying their severity according to the CEAP clinical staging system. This confirms that the system can assist diagnostic workflows with a high degree of reliability and interpretability. In addition to the statistical results, the analysis also confirmed that the model is ready to be potentially implemented in the clinic, especially in light of its low resource utilization and the conformance to ethical, sustainable and safety standards. The optimised training approaches used, such as early stopping, class balance, differential learning rates and staged data augmentation, also provided to its generalizability and use in a clinical setting. Overall, the results open avenues for further research and real-life application. The automated classification and staging system has great potential to support clinicians with timely diagnosis and treatment planning, thus providing better patient care while easing the burden on physicians. This is an important step forward in developing AI-supported, image-based diagnostic aides in vascular medicine.

CHAPTER 6

Conclusions and Future Scope

6.1 Conclusions

This research has successfully designed, implemented, and evaluated two high-performing deep learning models based on the ResNeXt architecture—one for the binary classification of varicose vein presence and another for multi-class CEAP severity staging. The integration of modern computer vision techniques with clinical problem-solving has resulted in a system capable of automating critical diagnostic tasks in the field of chronic venous disease assessment.

The classification model demonstrated strong performance, achieving an accuracy of 91.7% and an F1-score of 90.9%, which reflects a balanced and reliable capability in distinguishing between affected and unaffected cases. Similarly, the CEAP staging model achieved an accuracy of 89.3% and an F1-score of 88.6%, validating its ability to detect nuanced differences in disease severity—ranging from mild (C1) to severe (C6) stages—based solely on visual imaging features.

Key technical contributions of the project include:

- **Transfer Learning from Pretrained ResNeXt:** Leveraging pretrained ImageNet weights significantly accelerated convergence and boosted feature extraction from medical images, which typically suffer from limited data.

- **Fine-tuning Strategies:** Techniques such as layer freezing and differential learning rates allowed for optimal use of pretrained representations while still adapting to domain-specific patterns.

- **Early Stopping and Model Checkpointing:** These ensured that the models maintained generalization without overfitting to the training data, resulting in more reliable real-world performance.

- **Robust Validation Metrics:** By incorporating a comprehensive set of metrics including accuracy, precision, recall, F1-score, and ROC-AUC, the evaluation provided a well-rounded understanding of the models' clinical reliability.

The importance of this project exists not only in the rigorous clinical method applied, but also the accessible, pragmatic use it affords. The ability to detect varicose veins and clinically stage the level of severity using solely digital images represents an accessible and economical method that can support clinical decision-making, especially within limited-resource contexts. The value of a staging model to plan individual treatments is very useful in a progressive condition like chronic venous insufficiency (CVI).In summary, this project fully meets its objectives: to automate both detection and clinical staging of varicose veins using a deep learning framework for scalable, image-based, diagnostic capabilities within a health care setting.

6.2   Future Scope of Work

Despite the encouraging outcomes of the current study, there are several avenues for future research and improvement that can further refine and extend the impact of this work.

1. **Expansion of Dataset and Data Diversity**

   The dataset used in this study, though sufficient for prototyping, can be enhanced in the following ways:

   - Incorporating a larger number of images representing all CEAP stages equally to address class imbalance.

   - Including data from diverse demographics (e.g., age, ethnicity, lighting conditions) to improve generalizability.

   - Utilizing multi-center datasets to simulate variability across different clinical environments.

   These improvements would allow the models to perform better on unseen data and be more reflective of real-world scenarios.

2. **Incorporation of Multi-modal Inputs**

The current model relies solely on photographic images. Future work could consider:

- Integrating ultrasound or Doppler imaging to gain deeper insights into venous flow and anatomical structure.

- Including patient metadata such as pain levels, history, or duration of symptoms to contextualize predictions.

Combining these modalities would result in a more comprehensive and clinically grounded diagnostic tool.

3. **Advanced Architectures and Optimization**

While ResNeXt proved effective, further architectural improvements could be explored:

- Implementing Vision Transformers (ViTs) or EfficientNet for improved capture of long-range dependencies.

- Utilizing ensemble learning strategies to combine strengths of multiple models.

- Applying pruning and quantization techniques for model compression, enabling deployment on mobile or edge devices.

4. **Explainability and Clinical Trust**

Enhancing model transparency and interpretability is essential for clinical adoption:

- Using techniques such as Grad-CAM or SHAP to visualize and explain model decisions.

- Developing interpretability layers within the user interface to assist clinicians in understanding predictions.

5. **Deployment and Real-World Integration**

For practical impact, the model must transition from laboratory to clinical use:

- Creating a web or mobile-based application for seamless image upload and automatic diagnosis.

- Integrating with hospital information systems (HIS) for real-time image access and processing.

- Incorporating feedback loops where clinicians can validate or correct predictions for continuous model refinement.

6. **Post-Deployment Validation**

   Real-world effectiveness and safety must be established through:

   - Conducting clinical trials or pilot studies in hospital environments to assess usability and reliability.

   - Monitoring model performance for false positives/negatives in live settings.

   - Evaluating improvements in diagnostic speed, accuracy, patient outcomes, and cost-efficiency.

REFERENCES

[1] T. Sriranjani, K. P. Senthil Kumar and T. G. Thiruksha, "Varicose Vein Detection and Real-Time Integration using Faster R-CNN Algorithm," in *Proc. 3rd Int. Conf. Intell. Techn. Control, Optim. Signal Process. (INCOS)*, Chennai, India, 2024, pp. 1–6, doi: 10.1109/INCOS59338.2024.10527750.

[2] M. R. Thanka, S. J. Priya, E. B. Edwin, V. Ebenezer and R. P. Joy, "Varicose Veins Chronic Venous Diseases Image Classification Using Multidimensional Convolutional Neural Networks," in *Proc. 6th Int. Conf. Devices, Circuits and Syst. (ICDCS)*, Coimbatore, India, 2022, pp. 364–367, doi: 10.1109/ICDCS54290.2022.9780842.

[3] B. Oliveira et al., "Classification of Chronic Venous Disorders using an Ensemble Optimization of Convolutional Neural Networks," in *Proc. 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Glasgow, U.K., 2022, pp. 516–518, doi: 10.1109/EMBC48229.2022.9871502.

[4] E. Rabe et al., "Epidemiology of chronic venous disorders in geographically diverse populations: Results from the Vein Consult Program," *Int. Angiol.*, vol. 31, no. 2, pp. 105–115, 2012.

[5] S. Mirunalini, C. Jeyalakshmi and P. Muralikrishnan, "Fuzzy C-means based approach for analysis of Varicose Veins," *J. Pharm. Negat. Results*, pp. 2288–2295, 2022.

[6] A. S. Borde and G. V. Savrasov, "Mathematical modeling of varicose veins ultrasound heating," in *Proc. IEEE Int. Conf. Microw., Antennas, Commun. Electron. Syst. (COMCAS)*, 2019, pp. 1–5.

[7] K. Nithyakalyani and P. Muthu, "Detection of Chronic Venous Insufficiency Condition using Transfer Learning with Convolutional Neural Networks Based on Thermal Images," *Biomed. Eng. Appl. Basis Commun.*, vol. 35, no. 2, p. 2350030, 2023.

[8] H. A. Erdem, I. Erdem and S. Utku, "Designing a Low-Cost Early Diagnosis System Based on Deep Learning," *Int. J. Adv. Life Sci.*, vol. 15, no. 1–2, 2023.

[9] M. A. Das et al., "Early Detection and Prevention of Varicose Veins using Embedded Automation and Internet of Things," in *Proc. 5th Int. Conf. Inventive Res. Comput. Appl. (ICIRCA)*, 2023.

[10] K. Haritha et al., "Varicose Vein Diagnosis System and Therapy: A Review," in *Proc. Int. Conf. Comput., Power Commun. (ICCPC)*, IEEE, 2022.

[11] A. M. Abdalla et al., "Automatic Segmentation and Detection System for Varicocele Using Ultrasound Images," *Comput. Mater. Contin.*, vol. 72, no. 1, 2022.

[12] Roboflow, "Varicose Veins Detection Dataset," [Online]. Available: https://public.roboflow.com. Accessed: Apr. 2025.

[13] K. Kiranmai and K. B. Raja, "Deep Transfer Learning for Leaf Disease Classification," *Int. J. Intell. Syst. Appl.*, vol. 15, no. 3, pp. 45–56, 2023. [Online]. Available: https://www.researchgate.net/figure/ResNeXt-architecture-for-the-proposed-method-fig1-36012962

[14] K. Cherney, "Venous Insufficiency: Symptoms, Causes, and Treatment," *Verywell Health*, Jun. 22, 2023. [Online]. Available: https://www.verywellhealth.com/venuous-insufficiency-overview-4177929

[15] H. K. Lau et al., "Exploring Edge-Based Segmentation Towards Automated Skin Lesion Diagnosis," *Adv. Sci. Lett.*, vol. 24, no. 2, pp. 1095–1099, 2018.

[16] G. Mohammed, "Alzheimer's disease diagnostics by a 3D deeply supervised adaptable convolutional network," *Front. Biosci.*, vol. 23, no. 2, pp. 584–596, 2018.

[17] J. Kawahara et al., "BrainNetCNN: Convolutional Neural Networks for Brain Networks; Towards Predicting Neurodevelopment," *Neuroimage*, vol. 146, pp. 1038–1049, 2017.

[18] W. Shen et al., "Multi-scale Convolutional Neural Networks for Lung Nodule Classification," *Inf. Process. Med. Imaging*, vol. 24, pp. 588–599, 2015.

[19] Y. Yuan, M. Chao and Y. C. Lo, "Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks with Jaccard Distance," *IEEE Trans. Med. Imaging*, vol. 36, no. 9, pp. 1876–1886, 2017.

[20] M. A. Al-Antari et al., "A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification," *Int. J. Med. Inform.*, vol. 117, pp. 44–54, 2018.

[21] J. Xu et al., "Stacked Sparse Autoencoder (SSAE) for Nuclei Detection on Breast Cancer Histopathology Images," *IEEE Trans. Med. Imaging*, vol. 35, no. 1, pp. 119–130, 2016.

[22] J. Shi et al., "Ultrasound Image Based Tumor Classification via Deep Polynomial Network and Multiple Kernel Learning," *Curr. Med. Imaging Rev.*, vol. 14, no. 2, pp. 301–308, 2018.

[23] M. A. Moreno-Eutimio et al., "Enhanced healing and anti-inflammatory effects of a carbohydrate polymer with zinc oxide in patients with chronic venous leg ulcers: Preliminary results," *Arch. Med. Sci.*, vol. 14, no. 2, pp. 336–344, 2018.

[24] B. Kozma et al., "The Effects of Heat Exposure on Vaginal Smooth Muscle Cells: Elastin and Collagen Production," *Gynecol. Obstet. Investig.*, vol. 83, no. 3, pp. 247–251, 2018.

[25] P. Gloviczki et al., "The care of patients with varicose veins and associated chronic venous diseases: Clinical practice guidelines of the Society for Vascular Surgery and the American Venous Forum," *J. Vasc. Surg.*, vol. 53, no. 5, pp. 2S–48S, 2011.

[26] H. J. Welch et al., "Proposal for a national coverage determination for the treatment of varicose veins and venous disease due to disparate CMS local coverage determination policies," *J. Vasc. Surg. Venous Lymphat. Disord.*, vol. 5, no. 3, pp. 453–459, 2017.

[27] E. L. Chaikof et al., "The Society for Vascular Surgery practice guidelines on the care of patients with an abdominal aortic aneurysm," *J. Vasc. Surg.*, vol. 67, no. 1, pp. 2–77, 2018.

[28] B. Santler and T. Goerge, "Chronic venous insufficiency – A review of pathophysiology, diagnosis, and treatment," *J. Dtsch. Dermatol. Ges.*, vol. 15, no. 5, pp. 538–556, 2017.

[29] O. Nelzén and I. Fransson, "Varicose vein recurrence and patient satisfaction 10–14 years following combined superficial and perforator vein surgery: A prospective case study," *Eur. J. Vasc. Endovasc. Surg.*, vol. 46, no. 3, pp. 372–377, 2013.

[30] P. A. Patel et al., "Endovenous laser ablation therapy in children: Applications and outcomes," *Pediatr. Radiol.*, vol. 47, no. 10, pp. 1–11, 2017.

[31] Y. Liu et al., "Endovenous laser ablation versus conventional surgery for great saphenous varicose veins: Meta-analysis of randomized trials," *Zhonghua Yi Xue Za Zhi*, vol. 93, no. 23, pp. 1822–1826, 2013.

[32] G. D. Baxter et al., "Low-level laser therapy for the management of breast cancer-related lymphedema: A randomized controlled feasibility study," *Lasers Surg. Med.*, vol. 50, no. 1351, pp. 924–932, 2018.

[33] E. Depopas and M. Brown, "Varicose Veins and Lower Extremity Venous Insufficiency," *Semin. Interv. Radiol.*, vol. 35, no. 1, pp. 56–61, 2018.

[34] Y. Chen et al., "The directional migration and differentiation of mesenchymal stem cells toward vascular endothelial cells stimulated by biphasic calcium phosphate ceramic," *Regen. Biomater.*, vol. 5, no. 3, pp. 129–139, 2018.

[35] B. D. James and J. B. Allen, "Vascular Endothelial Cell Behavior in Complex Mechanical Microenvironments," *ACS Biomater. Sci. Eng.*, vol. 4, no. 11, pp. 3818–3842, 2018.