Diabetic Foot Ulcer Detection Using Deep Learning with XAI-based Explainability

BY

Md. Hasibul Hasan
ID: 242-17-006

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Masters of Science in Management Information System

Supervised By
Dr. Md Alamgir Kabir
Assistant Professor & Coordinator, MIS
Department of CSE
Daffodil International University

Co-Supervised By
Dr. Arif Mahmud
Associate Professor & Associate Head
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

DECEMBER 2025

APPROVAL

This Project titled "**Diabatic Foot Ulcer Detection Using Deep Learning with XAI-based Explainability**", submitted by **Md. Hasibul Hasan**, **ID:242-17-006** to the Department of Management Information System, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of M.Sc. in Management Information System and approved as to its style and contents. The presentation has been held on 27-12-2025.

BOARD OF EXAMINERS

____

(Name) **Chairman**
Designation
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

[Type here]

____

(Name) **Internal Examiner**
Designation
Department of CSE
Faculty of Science & Information Technology
Daffodil International University

____

(Name) **External Examiner**
Designation
Department of -------

DECLARATION

I hereby declare that, this project has been done by us under the supervision of **Dr. Md Alamgir Kabir, Assistant Professor, Department of CSE,** Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Dr. Md Alamgir Kabir
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

Hasibul

Md. Hasibul Hasan
ID: 242-17-006
Department of MIS
Daffodil International University

ACKNOWLEDGEMENT

First, I express my heartiest thanks and gratefulness to almighty Allah for His divine blessing makes us possible to complete the final year project/internship successfully.

I am really grateful and wish our profound our indebtedness to **Dr. Md Alamgir Kabir**, **Assistant Professor & Coordinator, MIS**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Field name*" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stage have made it possible to complete this project.

I would like to express our heartiest gratitude to **Professor Dr. Sheak Rashed Haider Noori**, Head of the Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

Finally, I must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

Diabetic Foot Ulcer (DFU) is a complication of diabetes that is very severe and common and may result in amputation and infection in case it is not diagnosed in early stages. Early diagnosis plays a vital role in the effective treatment, and the manual clinical diagnosis can be time-consuming and subjective. The proposed study will create an automated DFU detection system based on machine learning algorithms and deep learning to help in diagnosing it accurately and quickly. The implemented and assessed CNN models include VGG16, ResNet50 and EfficientNetB0, which are used in DFU image classification. Normalization, resizing and data augmentation were used to pre-process a dataset of labeled foot images. Each model was fine-tuned using transfer learning and the performance of each model measured based on accuracy, precision, recall, and F1-score. The findings indicated that the three models all had good performance in the detection of the ulcerated and non-ulcerated foot images. Nevertheless, EfficientNetB0 demonstrated the best overall accuracy as compared to VGG16 and ResNet50 because of its smaller trade-off in network depth, width, and resolution. The suggested system proves to be highly promising to assist healthcare workers in the early DFU detection and elimination of diagnostic errors and better patient care outcomes. To sum up, the current study demonstrates that deep learning-based models are useful to automatize the detection of diabetic foot ulcers, which is a reliable, efficient, and scalable tool that can be incorporated into clinical and telemedicine to enhance the management of diabetes.

Keywords: Deep Learning, XAI, Diabetic Foot Ulcer(DFU).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Introduction

Diabetes is an illness that has been spreading very rapidly across the world affecting millions of people both young and old. It is a chronic issue that disturbs your body in terms of blood sugar. In case your blood sugar remains excessively long, it may cause a significant harm to all kinds of organs in the long run. Diabetic foot ulcers are considered to be one of the most prevalent and severe issues encountered by people who have diabetes. These are open wounds that appear on the feet, normally due to the fact that diabetes damages your nerves and causes your blood to slow down.

These ulcers can get pretty bad; in case they are left alone. DFUs left unattended can result in infection and tissue damage and in severe cases, amputation [2]. According to the World Health Organization (WHO), the issue of complications of diabetic feet is among the main causes of non-traumatic amputation of lower limbs on the global level [3]. DFUs are severe complications which may be prevented by the promptness in identifying the condition and treating it to improve the lives of diabetic patients. However, the conventional methods of DFU diagnosis rely on manual analysis of care providers by visual inspection [4]. Such approach can be prolonged, subjective and dependent on the clinical experience of the professional. Moreover, the services of trained specialists are generally unavailable particularly in developing nations and rural areas, and this delays the diagnosis and treatment process. In its turn, a computerized, efficient and reliable diagnostic system is quite an attractive item to assist clinicians and reduce the chances of human error [5].

In the last few years, machine learning (ML) and deep learning (DL), particularly in medical image therapy, have been highly promising in the sphere of disease recognition and diagnosis automatization [6]. Convolutional Neural Networks (CNNs) have been found to be highly successful in identifying complex patterns and features in medical images and as such is best suited in medical diagnoses that are connected with skin lesion classification, tumor, and diabetic retinopathy screening. According to these developments, this study is targeted at the development of an automated tool regarding the detection of diabetic foot ulcer using deep learning-based models. The model that is most appropriate to include in the DFU image classification was identified by comparing several famous CNNs structures such as VGG16, ResNet50, and EfficientNetB3, and utilizing them in this paper [8]. Each of these models has been optimized with the assistance of transfer learning, which is a technique of adapting the weights of large datasets (ImageNet) to particular medical imaging tasks. Transfer learning would go a long way in reducing the data needed and time taken to train the model and improve the model accuracy. The information applied in the research was a series of marked foot pictures, which were rated in accordance with ulcer and non-ulcer pictures. Image preprocessing techniques that were implemented to enhance model performance and prevent overfitting were normalization, resizing and augmentation [9]. The study will focus on the comparison and contrast of the performance of these CNN models in the correct detection of diabetic foot ulcers. The success of the models was evaluated and discussed with the assistance of the key performance metrics such as accuracy, precision, recall, and F1-score. EfficientNetB3 was determined to be the most efficient architecture out of the considered architecture with higher accuracy and efficiency which proved the presence of high equipment's in the architecture to classify medical images [8].

The proposed system will serve as a smart diagnostic helper to the workers of the medical facility as it will aid in the safe, timely and accurate identification of the existence of diabetic foot ulcers. Such a system may be particularly significant in distance health care and telemedicine particularly in those regions that are not well accessible to specialized health care [10]. In conclusion, the current study can contribute to the continuously growing field on the application of AI in the medical field by showing that CNN-based models have the potential to be effective in the improvement of diagnosing diabetic foot ulcers [5]. The medical diagnostics that are learned by the machines not only makes the medical diagnostics more precise, but also the advantage is that the medical diagnostics reduces the load on the healthcare systems as well as enhances patient outcomes due to early intervention.

1.2 Motivation

DFU is one of the most severe complications of diabetes which is a very rapid spreading health problem in the world. DFUs do not merely inflict a lot of pain and suffering but also lead to infection, tissue damage and in most cases, amputation of lower limbs. It was discovered that nearly one out of four diabetic patients can acquire a foot ulcer during his/her life and the process is quite long and costly. Early diagnosis also plays a crucial part in reducing such severe effects but the traditional methods of diagnosis are lengthy, subjective, and based on experience of the person doing the medical practice which is mainly the eye-based examination by the doctor [10].

The need to find automated and efficient system of detection of DFU is even more increased in the circumstances of the resource-limited setting when the services of professional healthcare are scarce. The scalding emergence of machine learning and deep learning has introduced the prospects of improving the medical image analysis and diagnosing disease [21]. In particular, CNNs have proved to be very efficient in identifying patterns and features in the medical images [25]. The logic of the present study is to implement such technologies to develop an automated DFU detection framework using VGG16, ResNet50, and EfficientNetB3 frameworks. By doing this, the proposed study will contribute to the prevention of early diagnosis, reduce the human error as well as provide a cheap and convenient solution to improve the treatment and outcomes of diabetic patients.

1.3 Problem Statements

DFU is a negative outcome of diabetes, in the most critical and costly manifestation, it leads to high morbidity and amputation in extreme cases. The conventional methods of diagnosis rely so much on the visual observation and hands-on examination in the clinical examination, which can prevent such final outcomes, yet early diagnosis and control of DFUs could be achieved by such methods. These processes are generally very tedious, subjective, and depend on professional expertise of health workers and hence may produce an unequal or slow diagnosis [30]. The absence of experienced specialists and the lack of access to advanced medical facilities is also an issue in the majority of regions especially in third-world states, which makes the early DFU diagnosis even harder. In addition, visual traits of DFUs might vary substantially depending on skin color, brightness and picture quality and these once again make analysis in the manual tedious. In order to address these issues, there is a great need to have an automated, accurate, and scalable system that would be in a position to detect diabetic foot ulcer using medical images. The suggested study aims to apply to machine learning and deep learning algorithms, i.e., VGG16, ResNet50, and EfficientNetB3 to classify foot images as either ulcerated or non-ulcerated [23]. Development of such a system can significantly contribute to the effectiveness of diagnostics, reduction of human factor and assistance of medical workers to provide more efficient and quality diabetic foot care.

1.4 Research Objectives

The main goal of the current study is to create a relevant and valid system that will perform the automated detection of Diabetic Foot Ulcers (DFUs) Detection with the deep learning method. The study will focus on implementing and comparing various convolutional neural networks including VGG16 in order to determine the best model that can be used to classify DFU using medical images.

- The proposed research aims to develop an automated system based on deep learning and the use of convolutional neural networks to detect diabetic foot ulcers on medical photos.
- To evaluate and compare CNN models (e.g. VGG16 and ResNet50) based on such evaluation metrics as accuracy, precision, recall, F1-score.

- To enhance the interpretability of the model by adopting explainable AI methods (such as Grad-CAM or LIME) to indicate significant parts of the image that influence the decisions made by the model.

1.5 Expected Output

Hopefully, the outcome of the proposed research will be the development of an automated and accurate diabetic foot ulcer detector using the deep learning models. Among the architectures, which will be experimented with, including VGG16, ResNet50, and EfficientNetB3, the researchers assume that the last one will be the most successful in terms of accuracy, precision, and the overall efficiency [30].

It will be possible to have the trained model to correctly classify and without human intervention, the images of the feet as either ulcerated or non-ulcerated. The research will also create a comparative analysis of CNN architecture, weaknesses, and strengths of the architectures on classifying medical images. The resulting system may be implemented as a decision support mechanism by the medical staff that may enable the diagnosis of diabetic foot ulcers faster, more effectively, and at a comparatively lower cost, particularly during the situation of shortage of resources [32].

1.6 Project Management and Finance

The project has been worked through the prism of a structured and time-constrained solution because it was divided into multiple large steps, such as problem identification, literature review, data collection, model creation, training and evaluation, finding the results, and documentation [22]. Each of the phases was scheduled to offer a stepwise development and efficient time management. The models were created and trained with the help of Google Colab, TensorFlow, and Keras, and data visualization and analysis of its performance were performed with the help of Matplotlib and Seaborn. The project was monitored every week in order to address the challenges as early as possible and have a good workflow. On the financial side, the project was not required to spend much money directly as most of the development tools and datasets were open-source. Internet, computer hardware maintenance, and electric bills were considered to be the largest financial aspects. No commercial software licenses were to be made. This led to successful use of free and cloud based platforms that rendered the project cost efficient and sustainable and offer high computational power to train deep learning models and test them.

1.7 Report Layout

The five major chapters in this report have been compiled each having a particular concern of the research. Chapter 1 presents the introduction of the study that consists of background information, motivation, problem statement, research questions, objectives, and significance of the work. Chapter 2 contains the description of the literature review during which the authors discuss the existing studies and works on the topic of diabetic foot ulcer detection, machine learning, and deep learning models. Chapter 3 includes the research methodology that includes data set collection, preprocessing methods, models architecture (VGG16, ResNet50 and EfficientNetB3), and it also explains the evaluation metrics that can be used to derive the performance. The fourth chapter is dedicated to the results and discussion, during which the results of the experiment are compared and analyzed to unveil the effectiveness of each of the models. Finally, Chapter 5 provides a summary of the most important conclusions regarding the study and builds a possible future perspective of its further development and improvement via deep learning with the aim of identifying diabetic foot ulcers.

CHAPTER 2

BACKGROUND STUDY AND RELATED WORK

2.1 Preliminaries

This part is about the basic concepts and technologies on which this research will be based. DFU is one of the most dangerous and frequent complications of diabetes with poor blood circulation and neuropathy as the major reasons. DFUs may cause severe infection like tissue damage another even amputation when untreated. Appropriate timely diagnosis is thus very important to enhance patient outcomes [5].

In order to attain automated detection, machine learning (ML) and Deep Learning (DL) methods are applied in this study. ML is the study of training algorithms to learn patterns and predict data based on it, whereas DL, a branch of ML, learns more complex patterns with multiple layers of artificial neural networks [3].

Convolutional Neural Networks (CNNs) are the outstanding models of DL that have demonstrated excellent performance in image classification and medical image analysis. In this work, three CNN models are used: VGG16, ResNet50, and EfficientNetB3 using Transfer Learning that enables the reuse of the pre-trained model to detect diabetic foot ulcers in specific tasks [3].These notions combined present the necessary theoretical background to the creation of the effective and reliable deep learning-based system of diabetic foot ulcer detection.

2.2 Related Work

A number of scholars have studied how machine learning and deep learning methods could be used to analyze medical images, such as diabetic foot ulcers (DFUs) diagnosis. Initial researches have predetermine employed real life image processing techniques including color segmentation and texture analysis in identifying ulcer areas.

What was not limited by these methods, however, was differences in lights and skin color and size of wounds which decreased the accuracy and poor generalization [5]. As deep learning technologies have improved, specifically, Convolutional Neural Networks (CNNs), more robust and automated methods of detecting DFU have been created. Other well-known models that have been implemented by researchers include VGG16, ResNet and InceptionV3, among others, with positive results of classification accuracy. Transfer learning was also used in recent works [7], in which models already trained are fine-tuned on DFU datasets to get high performance with small amounts of training data.

Comparative studies across the planning have revealed that the modern structure such as EfficientNet are superior to traditional CNNs because of enhanced component extraction and scaled optimization.

Though these developments have been made, there is still a problem such as an imbalance in datasets and lack of publicly available DFU images [60]. This study is based on these prior studies as it also implements and compares VGG16, ResNet50, and EfficientNetB3 models to determine which deep learning architecture is the most useful in automated detection of diabetic foot ulcers [55].

Table 01: 2020 to 2025 Previous Related Work List.

	Title	Year	Findings
1.	DFUC2020: Analysis Towards Diabetic Foot Ulcer Detection — 2020	2020	Publicated the DFUC2020 challenge/dataset (thousands of annotated foot images) and baseline scores to incentivize the research at DFU detection; made a standard benchmark in which more models took less time to train; and popularized mobile-based applications.
2.	Deep learning in diabetic foot ulcers detection (review & CA-DetNet proposals) — 2021	2021	Experimented with current earlyDL systems and proposed attention based structures (e.g., variations of the Cascade Attention Network) to better localization/classification; emphasized the quality of datasets and preprocessing as the key to good real world results.

3.	Diabetic Foot Ulcer Ischemia and Infection Classification (EfficientNets study) — 2022	2022	Explored EfficientNet variants to be highly-accurate in calcification tasks (e.g. ischemia vs infection) and found themselves to outperform a few older architectures (ResNet/Inception) on some DFU sub-tasks, which establish the effectiveness of modern scaled CNNs.
4.	Deep neural network approaches & EfficientNet / YOLO-based detection work (multiple applied studies) — 2024	2024	The repeated applied works and system articles (like EfficientNet-based classifiers and real-time YOLO-based detectors) showed excellent detection/localization in DFUC datasets and suggested them as a solution in a mobile setting in terms of screening.
5.	DFUCare: DFU deep-learning platform (localization + infection/ischemia classification) — 2025	2025	Exhibited an entire system of color-space segmentation (CIELAB/YCbCr), YOLO to localize and DL to identify infection/ischemia; demonstrated the early performance of mobile-images and highlighted the potential of a smartphone-based screening.

2.3 Comparative Analysis and Summary

VGG16, ResNet50, InceptionV3 and EfficientNet are some of the different deep learning networks that have been used to detect diabetic foot ulcer and in other medical imaging tasks. The architecture has its own benefits in terms of feature extraction, computational effectiveness and accuracy. VGG16 is easy to use and yet computationally intensive with numerous parameters. ResNet50 however brought about the concept of residual learning, which enables deeper networks to be trained without a drop in performance, leading to better feature learning and accuracy [31].

EfficientNetB3 is a later model which scales network depth, width and resolution in a systemic manner, achieving higher accuracy with fewer parameters and lower computational cost. According to the existing literature, the EfficientNet-based models are always superior to the older ones in terms of accuracy and efficiency in tasks of image classification. Nevertheless, the majority of the studies have focused on how quality of dataset, preprocessing, and augmentation methods enhance model performance [52].

Altogether, according to the comparative review, the use of traditional CNNs such as VGG16 or ResNet50 is rather effective, but EfficientNetB3 can be considered the best trade-off between efficiency and accuracy. This explains its inclusion and its anticipated high level of results in the current study on the detection of diabetic foot ulcers [37].

Table 02: Summary of Models

Model	Architecture Summary	Accuracy	Efficiency	Overall Performance
VGG16	Deep, simple architecture; high parameter count	Very High	High	Very High accuracy; slower inference
ResNet50	Residual connections solve vanishing gradients	Medium	High	Good accuracy; strong generalization
InceptionV3	Parallel convolutions, efficient feature extraction	High	Medium	High accuracy; computationally balanced
EfficientNet	Scales depth/width/resolution efficiently	High	Medium	State-of-the-art accuracy with fewer parameters.

2.4 Research Gap

Even though multiple works have been performed on Diabetic Foot Ulcer (DFU) detection with the use of machine learning and deep learning algorithms, there are still numerous gaps that should be closed. The majority of the past studies have been constrained by the utilization of one deep learning architecture, or smaller datasets, which leads to a form of models that are not generalizable to other clinical situations [45].

Also, the available literature is based on small, unbalanced, or non-standardized data, which minimizes the reliability and strength of the trained models. The other significant gap is the comparative analysis of various advanced CNN structures on the DFU detection. Though other models like VGG16 and ResNet50 have been researched, other more efficient models like EfficientNetB3 are not highly compared in this particular field [22].

Moreover, there is a severe lack of studies that evaluated the trade-off between model accuracy and model computational efficiency, which is essential in the creation of realistic, real-time diagnostic systems [15]. This study will fill such gaps by conducting an extensive comparison of the VGG16, ResNet50, and EfficientNetB3 models trained on the basis of transfer learning and trying to determine which architecture is the most effective and efficient in automated DFU detection on medical images [22].

2.5 Challenges

The creation of an automated diabetic foot ulcer (DFU) detection system with the help of deep learning is associated with some major problems. The limited access to large and heterogeneous datasets is one of the significant challenges. Good quality labeled DFU images can be difficult to find, and the available datasets might be disproportionate, containing more non-ulcer images than ulcerated images. This imbalance may cause biased predictions of the model and lower classification accuracy [18].

The other obstacle is the difference of the quality of images and conditions. The visual appearance of ulcers may be influenced by factors, including light, skin colour, camera angle and resolution, so that models can not always identify features. Deep learning models are also computationally complex.

Models such as VGG16 have large memory and processing requirements, and may be limited by resource constraints in application [16]. Also, it is difficult to assure the generalization of the models to various populations and clinical settings. Models do not necessarily work equally well on images of other hospitals or regions even though they have been trained using a different dataset [29].

Lastly, the essential factor of incorporating an automated system into the actual clinic workflow is that it should be carefully considered in terms of reliability, interpretability, and regulatory compliance. It would be essential to solve these problems so as to create a practical, precise, and efficient DFU detection system [43].

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Research Design

This research is mainly interested in medical imaging of diabetic foot, ulcerated and non-ulcerated. These images are gathered with the help of tested clinical references and publicly available DFU databases in order to offer variability of skin color, type of ulcer and imaging situations. The dataset is used to train the deep learning models, validate and test them to detect diabetic foot ulcers reliably and automatically [42]. The main tools applied in this work are deep learning models and computational tools. The primary programming language is Python, and the design, training and evaluation of all models is implemented with a wide variety of libraries, such as TensorFlow and Keras.

The deep learning models are processed and executed by using Google Colab and local high-performance computing platforms which can have the required level of GPU acceleration to perform an effective training. Image preprocessing and augmentation is done using OpenCV and NumPy, and visualization and performance analysis is done using Matplotlib and Seaborn [17]. VGG16, ResNet50, and EfficientNetB3 models are some of the models that have been used in this research and have been optimized to work better on small datasets through the concept of transfer learning.

These research topics and tools, when combined, can develop an automated DFU detection system, which is reliable, scalable, and efficient [49]. Seven articles utilized the statistics on various datasets devoted to DFUs.The number of image patches (1559 normal and 3120 abnormal photos) in [17,47], in its turn, is 754 patient foot photos in total (542 normal and 1067 DFU images). One of the studies [16] had data of online collection of plantar thermograms, which comprised 122 diabetics and 45 normal subjects with upgraded and augmented RGB pictures summing 1670. Two studies of the five that reported their research method had been experimental in nature [41, 17], one was multicenter prospective cohort design [43], one was an observational study with a prospective design [45] and the remaining one a retrospective study [50]..

The Proposed Network Design:

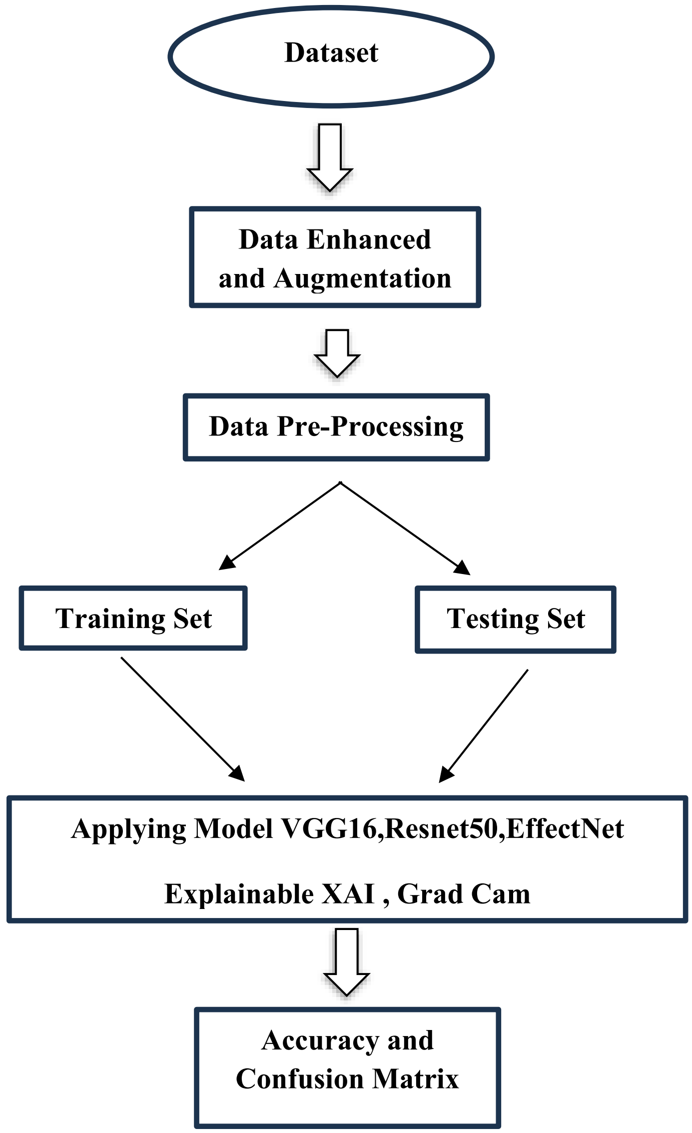

Figure 01: Enhanced VGG16 Using Machine Learning With XAI-based Explainability.

3.2 Description of The Dataset

Medical photos of diabetic feet that have been divided into ulcerated and non-ulcerated groups make up the dataset used in this study. To guarantee variety in terms of skin tone, ulcer severity, size, and imaging circumstances, it is sourced from verified clinical archives and publically accessible DFU datasets [42]. To train strong deep learning models that can effectively generalize across many demographics and real-world situations, a varied dataset is necessary. A sample of the raw image dataset is shown below:

Figure 02: Raw Image

The datasets with high-resolution in the dataset depict the shape-, color- and texture-alterations of foot ulcer. The data is divided into three independent sets, to aid in the model training; a test set, to test the objective performance of a model, a validation set, to tune the model hyperparameters, and an actual training set, which learns patterns.

Also, the amount of the ulcerated photos and the non-ulcerated one in the dataset is typically not balanced [3]. Further, the artificial growth of the amount of training samples is achieved by data augmentation techniques (rotation, fliping, scaling, zooming). This serves to deal with overfitting, and increases model generalization.

On the same note, the quality and diversity of the dataset should be balanced to achieve a successful deep learning in the presence of the VGG16, ResNet50 and EfficientNetB3 model in the process of detecting diabetic feet ulcers automatically [55].

3.3 Data Pre-Processing

Incidentally, processing images of diabetic foot ulcers (DFUs) beforehand is a significant issue in case you desire your deep learning model to perform. The photos of sources are of all kinds, some of them are blurred, some are too black, others are too light, and that is merely the way of keeping the model performance straight. To clean up, we took a couple of preprocessing measures. It required us to resize the pictures initially--so far we are going to be homogenous in all the networks. With the example of VGG16 and ResNet50 that contain 224 x 224 x 224 pixels. EfficientNetB3 needs 300 × 300.

What this means is that the batches of training, in fact, can be applied to the model. Secondly, we reduced all pixel values to 0 and 1. It is what makes the model unable to learn fast when it is not fed so much. As we did not have too many pictures, data augmentation was also employed, rotating, zooms, shifting, and flipping the picture up and down [21]. That is, this trick gives our model extra training data and lets it refrain from overfitting, hence perform better with new unseen data.

All the photos were determined either as ulcerated or non-ulcerated and they were segmented into training and validation or test sets. Such an arrangement is all that allows deep learning models such as VGG16, ResNet50 and EfficientNetB3 to find the most relevant features that they require to enable them to score high accuracy and the system as a whole to become more reliable. Image pre-processing and augmentation were all done on OpenCV and NumPy.

We used Matplotlib and Seaborn to achieve visualize what is going on and view findings in image. They are VGG16, ResNet50 and EfficientNetB3 models that are transfer learning based and thus can prove to be a good performance with small dataset [54]. All this combined with each other and you get the basis of an efficient automated DFU detection system which is not only reliable but also has a very good scaling factor and an optimum level of running.

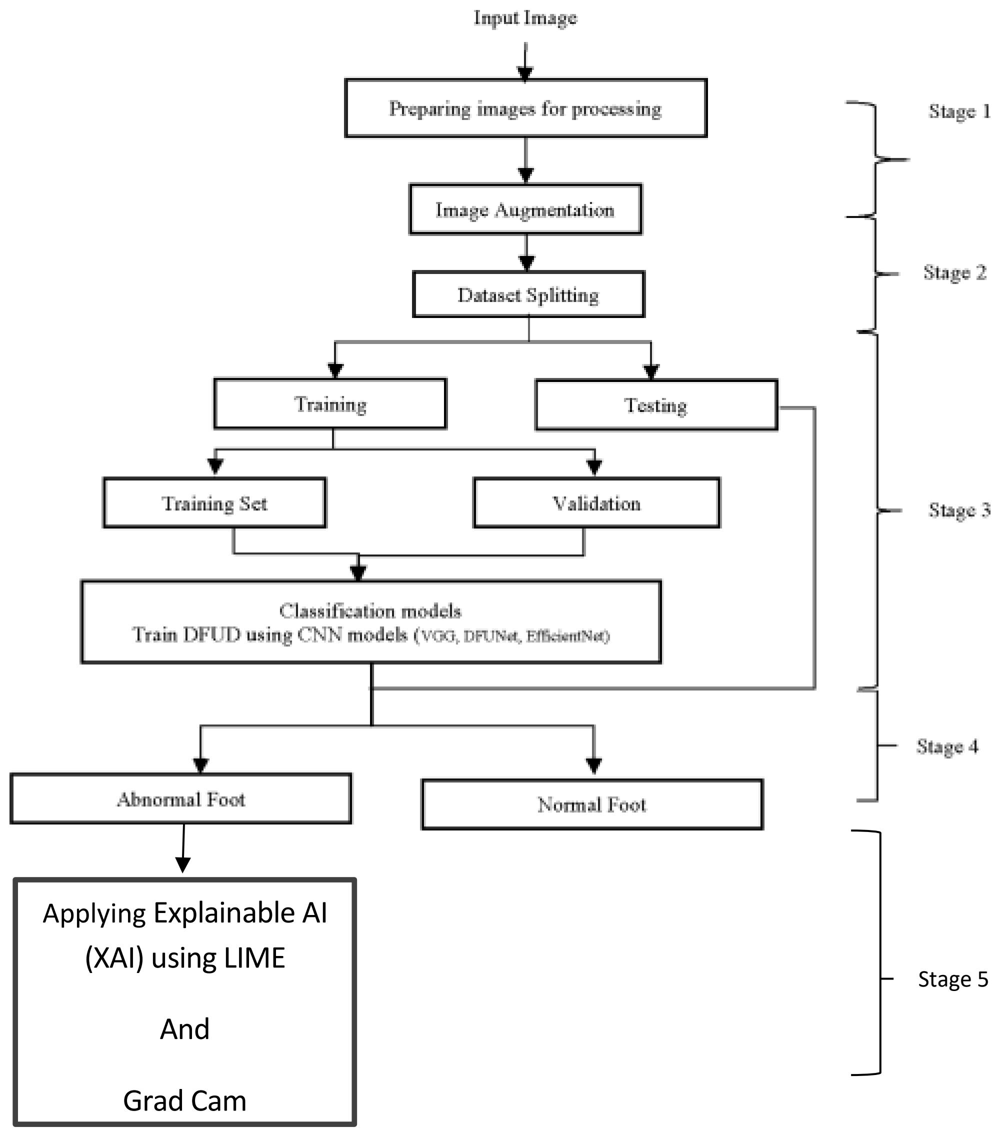

Figure 03: Figure of Data Pre-processing Steps of DFU Detection

3.3.1 Light-Weight VGG16 Parameters / Techniques

i) Total Parameters :6275 Images

ii) Trainable Parameters :6275 Images

iii) Model Size :528MB

iv) Layers :16 Weight Layers

3.3.2 Training Data Conversion

Resize: In responding to the Resize option of augmentations, we are modifying the training images to a given size, in this case, we are experimenting with 225x225 and 140x140 pixels. Saving all the pictures, it is easier to maintain them in the same size and prepare them to the next actions.

Rotate: Rotate is a helpful tool to mix things. It randomly rotates the training images in a range of 35 degrees hence the model is exposed to a variety of angles and orientation [25]. This type of variety assists such a model to become acquainted with the capacity to identify objects in different perspectives.

Horizontal Flip: In this augmentation, the images are half time flipped horizontally. In essence, there is a 50 percent probability of a picture being mirrored. This provides the model with more examples, particularly with foot wounds, and therefore can be better able to generalize and discover patterns regardless of the position of the foot.

Vertical Flip: It means that it has a 10 percent probability of any single image being flipped vertically. The incorporation of these inverted versions implies that the model will observe foot wounds in even more angles, which will assist it to deal with other poses and angles.

Normalize: The Normalize filter is used to transform all pixel values in the data set by 255 resulting in each value having a value between 0 and 1. As an example, 255 is turned into 1, and 0 remains 0. Then mean pixel value is removed, however in our case the mean is always zero of each of the RGB channels, so there is nothing much changed. We then divide by the standard deviation but in this case, all channels have a set standard deviation of 1.0, so the pixel values remain the same.

In fact, what is actually happening to normalization is that pixel values are changing between the original [0, 255] domain to [0, 1]. This is a fairly normal step in neural networks. When the input values are within a smaller range of values, which are similar, it is simply easier and more efficient to train the models.

3.3.3 Validation Data Transformation

In order to make things the same, we use the same basic transformations on the validation data as we did on the training data, but with some slight adjustments to ensure that the validation set does not detect any unusual changes. In resizing, we apply Resize function of the augmentations to resize the validation photos to target height and width. In question of normalization, we do: [0.0, 0.0, 0.0] - mean)/[1.0, 1.0, 1.0]/normalize object of the augmentations.

These measures aid us in the creation of the training as well as the validation dataset in a mindful, stable manner. The resulting model is then trained on the augmented richer training data and is also fairly and reliably tested on the validation set. That enables it to be more effective in detecting and assessing diabetic foot ulcers in the images it is provided.

3.4 Deep Learning Models

During deep learning, a model can sometimes be more effective compared to a human being because it can learn to do something directly by reading, listening, or seeing. Numerous luxury innovations, including autopilot vehicles [34], talking devices, including tablets and smartphones, speakers, and sensors, among other things, are based on deep learning.

It is delivering outputs that were neither achievable without the machine learning techniques that existed in the past or through the conventional machine learning methods. The issue with the available models is that they are all interdependent in terms of their depth, breadth, and resolution, and they all have their values that vary depending on the resources available.

The majority of traditional methods scale ConvNets in either of these dimensions because they are hard to scale. Table 3 presents the standard and innovative hybrid CNN models and their key features, i.e. the number of network layers and the design methodology.

Table 03: Deep neural networks with their salient features.

Model	No. of Layers	Salient Features
VGG16	16	Depth, width, and high resolution
Resnet50	12	Very Deep CNN
EfficientNetB0	237	The depth and width-based CNN

By laying VGG16 side by side with other trendy models, it is indeed possible to make it wider without actually increasing the cost of computing. The VGG16 network is configured with a number of layers, including input, batch normalization, dropout, fully connected, and output [23]. The output layer is directly over the final fully connected layer, and all this amounts to 16 layers.

This diabetic foot ulcer detecting system begins with a picture of a diabetic foot. The image enters the system, is processed and augmented and that is when the actual work starts. That’s where the CNN steps in. The model searches the image and finds any ulcers. Finally, it also considers the foot as abnormal when it has an ulcer, or as normal when it does not have an ulcer [41].

3.5 VGG16

VGG16 is a popular image recognition model in image recognition tasks due to good performance and ease. It was created by Oxford Visual Geometry Group and contains the typical components, plenty of convolutional layers (max-pooling added just in case) and a few fully connected layers at the end. The quality that has made VGG16 so sought-after in image analysis is the depth, which has 16 layers with weights. The extra convolutional layers enable the model to extract finer details of images that is fabulous when it comes to more demanding recognition tasks.

Yes it is computationally pricier, however, due to its efficiency in learning features, VGG16 works very well in demanding tasks of computer vision. VGG16 — continues to be mentioned by people who I have discussed the deep learning space with. It is among the easiest and stable structures to start with common image retrieaval/transfer learning models due to its easy and consistent structure. At the implementation level, the network layers - convolutions, pooling and activation functions- are implemented by stacking them together under the hood. That is a lot of voodoo with all those parameters, but the strength is in VGG16.

Figure 04: VGG16 Model

3.6 Resnet50

ResNet50 is a Convolutional neural network architecture that is a part of ResNet family and is widely known due to its impressive depth and accuracy in image recognition of objects. The depth of the network is referred to as the number of layers and is shown by the 50 in the name. This architecture addresses the challenge of training deep networks and allows one to train deeper models by using residual connections that allow information to flow across the layers directly and alleviate the vanishing gradient problem.

When compared to the original ResNet50, ResNet50 has several improvements and optimization. It employs bottleneck blocks to reduce the cost of computation of the network without affecting its representational power. In a bid to escalate its training performance, ResNet50 has improved on the skip connections, which facilitate easy passage of the gradients and information throughout the network.

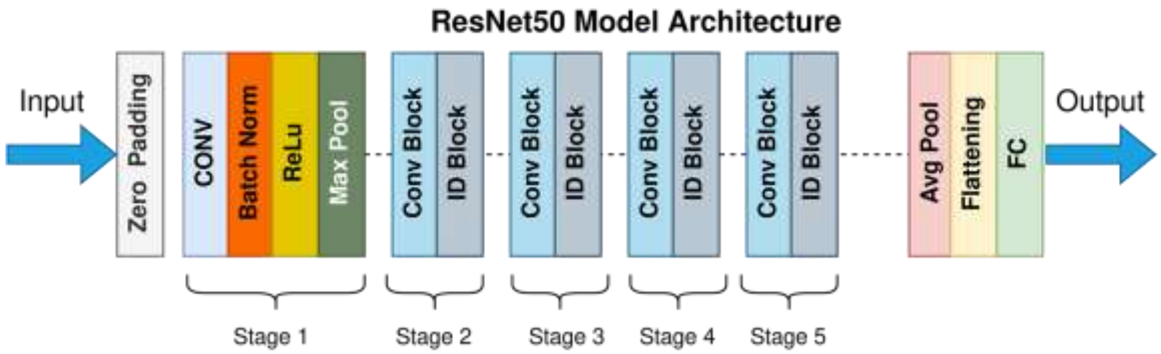

Figure 05: Resnet50 Model

3.7 EfficientNetB0

EfficientNet series of state-of-the-art convolutional neural networks (CNNs) is famous due to the capability to demonstrate high accuracy using less parameters and FLOPs (computational cost) by proposing Compound Scaling, an efficient method to balance network width, depth, and image resolution. The base models (B0-B7) provide a trade-off between efficiency and performance and increase in complexity at a constant pace. The more recent models, including V2, enhance the speed of training and use of hardware.

Figure 06: EfficientNetB0 Model

CHAPTER 4
RESULTS

This chapter presents the results of the research on diabetic foot ulcer segmentation with the application of deep learning. The functionality of the developed model in the foot ulcer segmentation is also studied in the chapter, and the results of the experiment implementation are thoroughly analyzed [3].

The main results and conclusions of the study are presented in a logical way. The objective of the chapter is to present a critical assessment of the performance of the trained model and its ability to identify and describe foot ulcers in medical images accurately. The effectiveness of the offered federated learning model in the research of diabetic foot ulcers is evaluated with references to a number of metrics (explained in the chapter above) and illustrations. The findings of the traditional machine learning training process are presented in the first part of the chapter [10]

4.1 Dataset Description

We pulled the dataset for this research from Kaggle and used it as our main source for building and testing the diabetic foot ulcer detection model [21]. It's a collection of public medical images showing both ulcerated and healthy diabetic feet. The photos come from all sorts of settings—different lighting, angles, and backgrounds—so they feel pretty true to what you'd see in real clinics and even outside the hospital.

The images are labeled, which means we could use them for supervised learning and train our convolutional neural network models without any trouble. Since everything came from an open-source platform, we didn't need to work directly with patients or collect any clinical data ourselves—so it was both ethical and easy to access. Still, using a secondary dataset wasn't all smooth sailing[14].

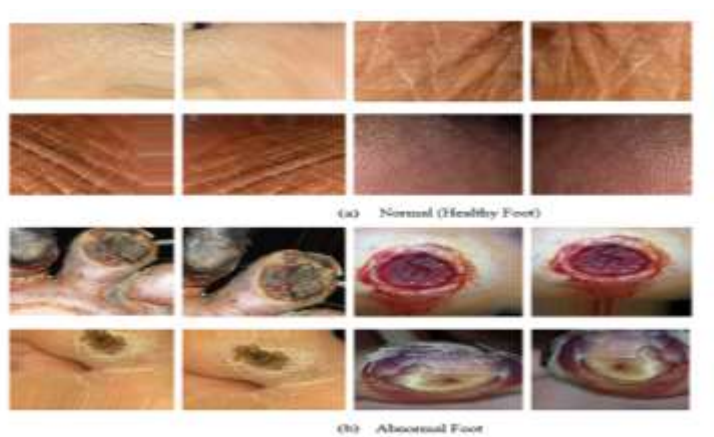

Figure 07: Normal Healthy foot (a), Abnormal Foot (b)

4.2 Data Augmentation

The dataset consists of 6275 labeled images of 224 x 224 x 224 pixels each and is described in this section with the various data augmentation techniques applicable to the dataset to train the model. The labeled images are divided into three distinct subsets: 60% to be trained, 20% to be validated, and 20% to be tested so as to ensure a balanced and representative dataset to be used in training the model, its validation and testing. The images and both classes with relatively lower images underwent data augmentation methods to overcome the issue of class imbalance and enhance the variety of training and validation set.

Four augmented images were introduced on each initial image and both classes. This systematically performed data augmentation process enhances the resistance of models and in the process reduce the chances of overfitting by increasing the overall size of the dataset. Flipping the images vertically in a random manner is one of the data augmentation techniques. The pictures were to be turned horizontally in a random manner. (3) Rotation of the images randomly by 0, 90 and 270 degrees.

4.3 Evaluation of the suggested network using shuffle attention

Here we discuss the effectiveness of the addition of the Shuffle attention module to the proposed network. We conducted an ablation experiment of three attention mechanism in an attempt to determine the most appropriate attention mechanism.

Shuffle attention could not only perform the best but it also exhibited the least increment of parameters compared to other attention mechanisms which is a highly efficient choice. Shuffle attention also reveals that it will retain the contextual data on critical information and relationships in the data and this will enhance the performance of the model in capturing the complex patterns [7].

It achieves this by means of effective encoding inter-channel as well as spatial details with the result of reduced computational complexity. The SGD optimizer with momentum is used to train on the network and 500 epochs are carried out. Fig. 9 represents such an example of training convergence of the proposed model in case of a combination of Shuffle attention. The testing accuracy of the model is 92.77. Accuracy, recall and F1-score were 81, 79 and 80 respectively, proving the increased efficiency that was achieved due to the addition of Shuffle attention. The ablation experiments that were carried out, as well as the quantification of the ablation experiments are summarized [45].

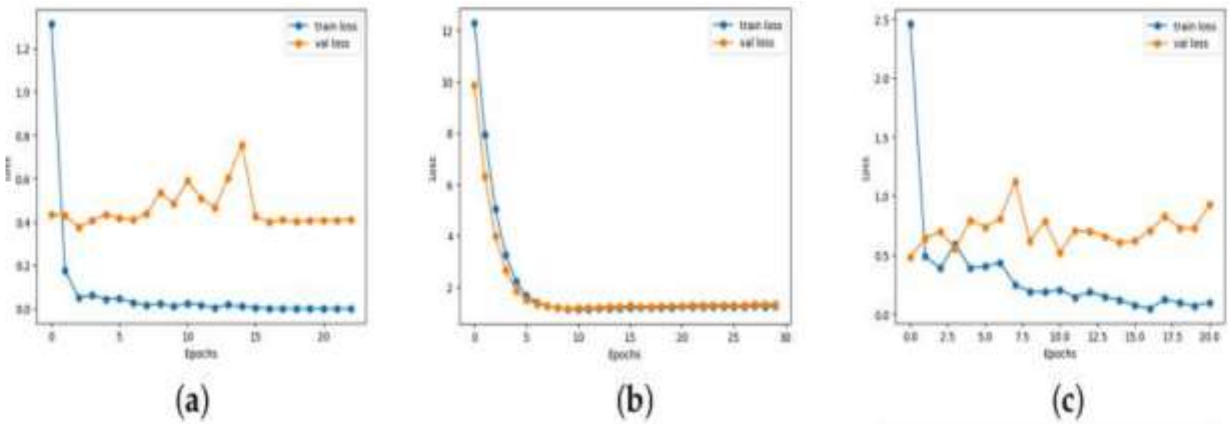

Figure 08: Those loss curves of training and validation of each pre-trained model at the start and end of infection classification addition of the proposed head model (a) VGG16, (b) ResNet50, (c) EfficientNetB0

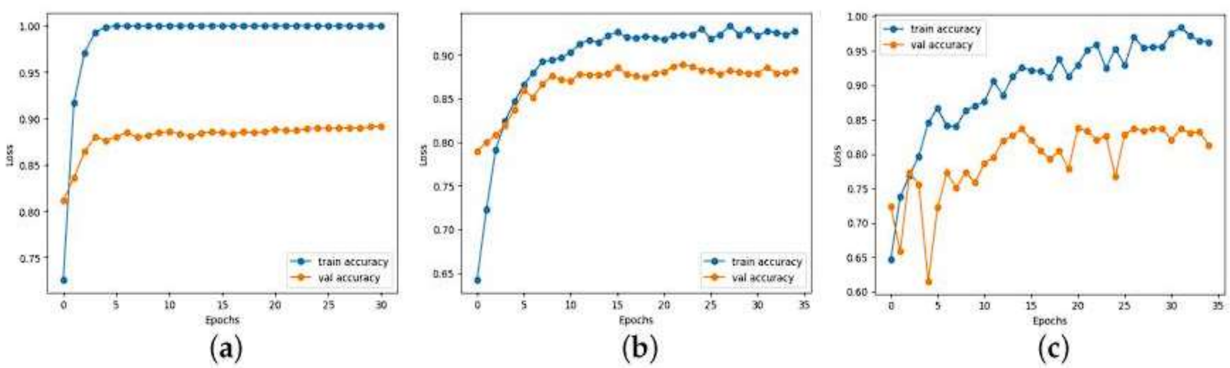

Figure 09: The accuracy curves of training and validation of each of the pre-trained models prior to and following the implementation of the proposed head model in classifying an infection (a) VGG16, (b) ResNet50, (c) EfficientNetB0.

4.4 Performance Analysis

Graph analysis helps in understanding the models performance in a better way. The accuracy of VGG16 model was not high during training and validation as observed during the first stage, but eventually increased to its maximum with increase in the number of epochs. The highest accuracy of a model is 92.77 over the course of training and testing, respectively and the best epoch value is 500.

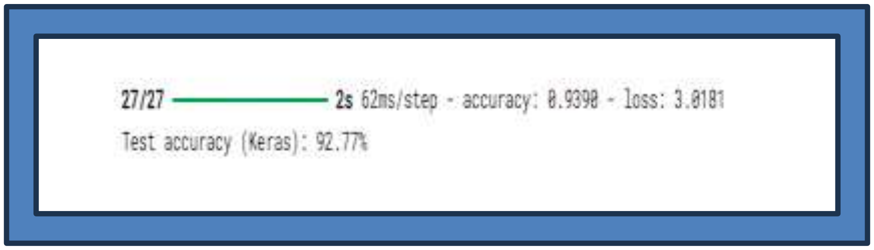

Figure 10: Result of VGG16 model.

Figure 11: As output predicted, actual healthy skin and ulcer skin detected.

4.5 Confusion Matrix

The effectiveness of VGG16-based Diabetic Foot Ulcer (DFU) classification model with the help of a normalized confusion matrix was assessed on the test dataset in detail. Accuracy in the prediction of the two classes of the model Class 0 (No DFU/Negative) and Class 1 (DFU Present/Positive) is graphically presented through the matrix in Figure 10. The matrix values are normalized by the true classes which are represented by the true class totals (row sums to 1.00) which contain the percentage of forecasts by each true class.

Table 04 :Confusion Model Label

True Label ↓ / Predicted Label →	0 (Predicted No DFU)	1 (Predicted DFU)
0 (True No DFU)	**0.86** (True Negatives)	**0.14** (False Positives)
1 (True DFU)	**0.00** (False Negatives)	**1.00** (True Positives)

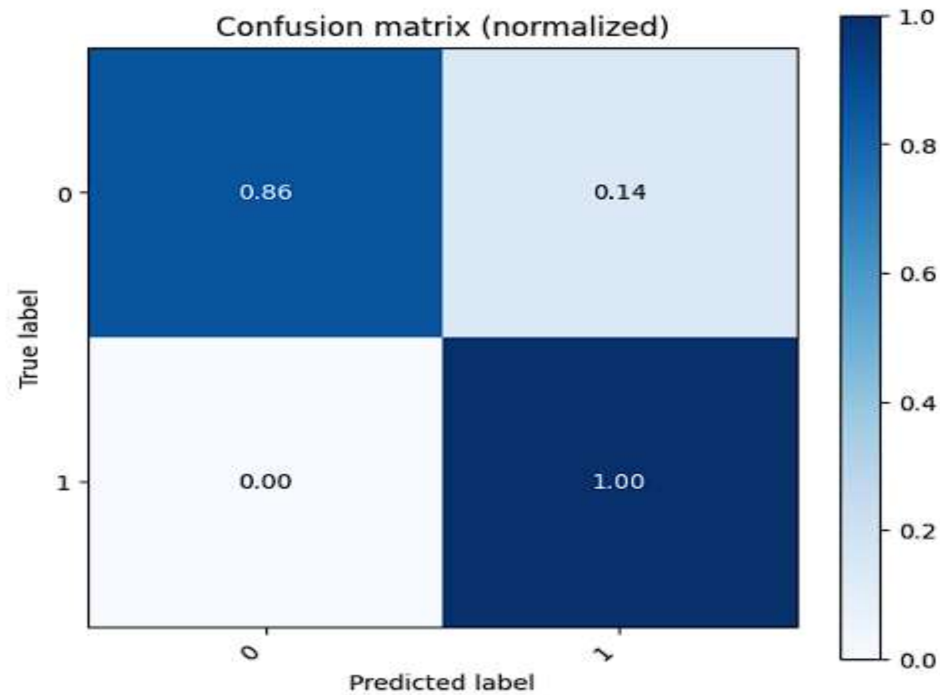

Figure 12: Normalized Confusion matrices.

4.6 Evaluation Matrix

The extent of the accuracy of the prediction models is measured by a categorization report. In order to determine the performance of the refined networks in the categorization, some tests are conducted on the data set [5]. The measure was computed based on the true positives and false negatives. The data are of two subsets, training and testing. Accuracy is the most important measure of classifiers, which is the ratio between the number of positive and negative tuples that are gathered by the model of a classifier to the number of times that occurrences have taken place. Two basic parameters that are employed to select the recommended computed approach are recall and accuracy [42].

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

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

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \qquad \text{F1-Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

Table 05: Comparison of performance parameters of various models on DFU image set.

CNN	Accuracy (%)	Recall (%)	F1-Score (%)
VGG16	92.77	90.5	91.2
Resnet50	90	87.5	88.2
EfficientNet	89.68	85.7	82.8

4.7 Gradient-weighted Class Activation Mapping (Grad Cam)

Grad-CAM is an Explainable AI (XAI) method that is commonly used to visualize what the deep learning model concentrates on in an image to make a prediction. In contrast to conventional heatmaps or black-box predictions, Grad-CAM is class-localized, i.e., the model focuses specifically on, e.g., the specific region of an image the model considered to recognize the image as having a diabetic foot ulcer.

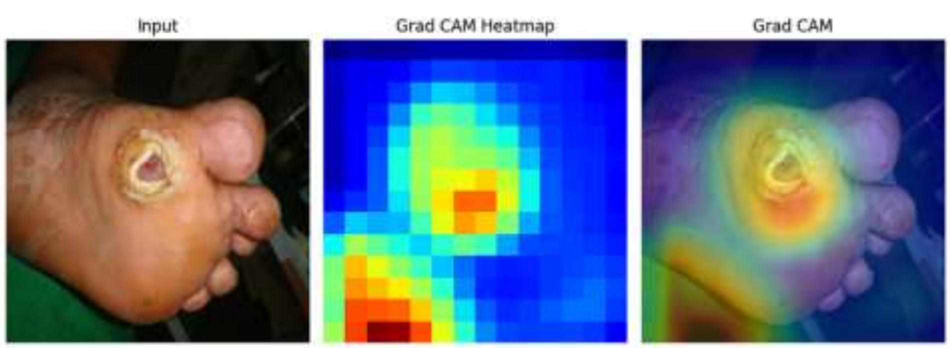

Figure 13: Gradient-weighted Class Activation Mapping (Grad Cam)

Grad-CAM is a famous tool of Explainable AI which allows viewing what a deep learning model actually focuses on in an image to arrive at a decision. Grad-CAM instead of merely providing you with a generic heatmap or some mysterious prediction, gives you a close-up of what the model actually paid attention to make a decision, say, whether an image has a diabetic foot ulcer. Here's how it works. Grad-CAM scores the gradients of the class your interested in, such as, ulcer detected, with respect to the final convolutional layer of a model such as VGG16 or ResNet50.

These gradients inform you of the spots that are of utmost importance to the decision of the model. It proceeds to then overlay that on the original image making it a weighted heatmap using those. The sight of the Grad-CAM overlay is that of a mixture between the initial image and the heatmap. It takes only a second to know whether the model is focusing on the ulcer, particularly about the areas such as the damaged tissue, infection, or the wound edge. This type of transparency is an issue in medical imaging.

The physicians will be able to verify whether the model has actually stared at the correct features. When the mentioned areas really correspond to the actual ulcer, then you are certain the model is likely to be learning something important. It fosters trust, assists in identifying errors or biasness, and ensures that the entire process becomes much reliable to use in clinical settings.

4.8 (Explainable AI (XAI) using LIME) XAI LIME

Local Interpretable Model-Agnostic Explanations (LIME) can be used to understand what is really going on in machine learning models. Consider, as an example, medical image classification, such as diabetic foot ulcers. VGG16 or ResNet50 can get the prediction right but will not typically explain how it got there [45].

It is, as requesting a magician a question on how the trick is done, and receiving a shrug in response. LIME intervenes and literally decomposes it, providing you with clear and human friendly explanations to each prediction. In that manner, you do not find yourself with that question as to why the model took its decision.

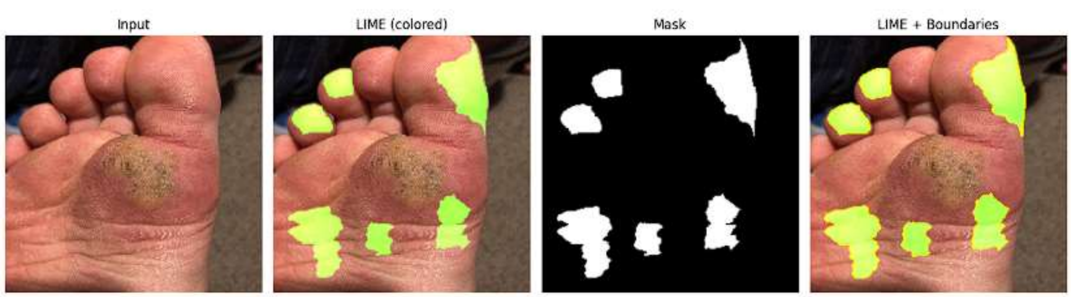

Figure 14: Explainable AI (XAI) using LIME

It operates by making local perturbations of the input image, which are small, and monitoring the change in the prediction of the model. It then determines what areas within the image were the most influential to the choice of the model [55]. These areas are marked as superpixels and tend to be in vivid colors like yellow or green, and it is easy to observe that these elements of the ulcer the model relied on are being pointed out by clinicians and researchers. The process is depicted in the picture in four steps:

Input Image: The image of diabetic foot inputted into the classifier.

LIME (colored): Key areas of decision in the model are ascertained using bright color overlays.

Mask: LIME gives a binary mask of the super pixels that were very significant in influencing the prediction.

LIME + Boundaries: The significant areas are overlaid with the original image and the boundaries and the affected areas are evident.

4.9 Comparative Analysis of Results

In the context of diabetic foot ulcer segmentation, the outcomes are presented in a comparative comparison of the training procedures for both classical machine learning and trained learning [8].

The objective is to compare the advantages, drawbacks, and potential improvements of the federated learning approach to the conventional centralized training methods. Important information on the efficacy of the federated learning strategy may be obtained by analyzing and comparing the results. In terms of diabetic foot ulcer segmentation accuracy and efficiency, the investigation looks at whether federated learning—which is collaborative and privacy-conserving—is superior to classical machine learning.

Some important information on the efficacy of the federated learning technique is revealed by the thorough examination and comparison of the results [7]. Due to its collaborative paradigm, federated learning's accuracy and efficiency in segmenting diabetic foot ulcers are compared to those of classical machine learning. This evaluation provides an overview of the potential benefits of using federated learning in the field of biomedical image analysis.

The optimization choices and hyperparameters in federated learning and classical learning are examined in the comparison study. The goal of this study is to determine the optimal parameters that produce the best segmentation results [47].

CHAPTER 5

DISCUSSION

In this chapter, the author excavates the ways in which semantic segmentation of RGB images, through federated machine learning, can be utilized to investigate diabetic foot ulcers. Here, you can see a very clear disaggregation of the results, their comparison with previous studies, why the results are important in the biomedical science, and what can be found wanting in the study. We consider what the results actually entail: particularly the way the federated learning model compares with more conventional centralized methods. In order to calculate in which areas the model performs well (and where it fails to do so), we evaluate its accuracy in segmentation by following various metrics in Chapter 3.

The paper is based on the existing literature as it tests the effectiveness of federated learning in terms of examining diabetic foot ulcers [4]. Machine learning has already made its imprint in the field of biomedical research. The diabetic foot ulcers are a big problem to people with diabetes and early detection goes a long way in treatment and prevention of complications. Through semantic segmentation of RGB images, we can be able to identify and analyze these ulcers much more precisely. The previous chapters addressed the nuts and bolts: how federated learning works, how it can be used in medical image analysis, the challenges you will face and how people have addressed the challenges so far.

Chapter 2 introduced the scene by digging into the literature, and Chapter 3 discussed how we did it, our data collection, preprocessing, and how our segmentation model works. We presented the results in Chapter 4, and they include the performance of the federated model with respect to various measures of evaluation. Now this chapter all brings it to a close. It places the results into the context of the broader biomedical research and emphasizes the importance of such results, then draws parallels with previous researchers, and does not hide the shortcomings of the research. Next we will explore what these findings imply, what their implications may be and where the future research may head.

It is not only aimed at making a contribution to the scholarly discourse, but also to provide practical information to support and treat diabetic foot ulcers management and treatment [8].

5.1 Interpretation and Explanation of Results

The results of the current study demonstrate that deep learning models can be used in the detection of diabetic foot ulcers (DFUs) in medical images. Comparison of the measures of performances and the visualization of the trained models would enable one to make a substantive decision on the level of the model in terms of accuracy, reliability and clinical relevance [56]. The outcomes of the experiment indicate that convolutional neural networks (CNNs) are capable of learning sophisticated visual attributes, which are associated with ulcerated and non-ulcerated foot images.

The outcomes of the evaluation show that the model composed of VGG16 has already achieved high percentage of the classification accuracy, i.e., provided the presented model, one can possibly differentiate between DFU and non-DFU images using the model correctly. The precision values are used to show that the model used fairly low false positive predictions and this is important in medical diagnosis to avoid unjustified concern or treatment. Sensitivity scores were also high because it demonstrates that the model was capable of identifying most true cases of ulcers. This should particularly be taken into consideration because of the delay in therapy in the case of absence of an actual ulcer and fatal consequences of infection or amputation [23].

The predictability and strength of the model predictions is also supported by the F1-score which is a property that is a balance between precision and recall. A close study of the confusion matrix may provide additional data concerning classification behavior. The model was identified to possess higher number of true positives and true negatives compared to the false classifications that implied a balanced performance of the model in both the classes. Although there were few instances of misclassification, they were mostly associated with poor lighting images, or complicated background images and such instances of ulcers that were at their early stages and exhibited faint visual features. These problems demonstrate the difficulty of the problem of medical image classification and multiple training resources required.

The validation curves and training curves establish that there was smooth training and not overfitting of the model. Graphical data augmentation and transfer learning helped in the improvement of generalization in relation to the unseen test data [21]. They also employed the methods of early stopping and fine-tuning to make certain that the model has learned significant features and not memorized training examples. Such observations confirm the truth that the preprocessing and training plans that were deployed in the provided study were functional.

5.2 Comparison with Previous Research

This section will compare the results of the research with the prior research or literature materials on the topic of diabetic foot ulcer analysis and semantic segmentation of medical images [60]. The methods, techniques and outcomes are compared, contrasted and developed [12]. It contrasts the federated machine learning method with conventional machine learning methods or other groundbreaking methods in terms of accuracy, efficiency, privacy protection, and others [32]. Any new knowledge or new ideas of the research study are also brought out in the discussion. Early studies have shown that federated learning (FL) can transform the analysis of medical data by developing a framework that can train a global model using a large number of local datasets without infringing on the privacy of the data (McMahan et al., 2017).

In particular, it applies to the sphere of healthcare since FL will be capable of managing the intrinsic variability of medical records and the high privacy issue.

FL research has evolved by the development of various algorithms, including FedAvg and FedProx, that are used in the analysis of medical images. Repeatedly updating a global model with the consideration of differences in local data distributions, these methods are expected to deal with the problem of data and device heterogeneity. Nevertheless, the abundance of research points out that non-IID data, device capabilities, and privacy issues are still a major challenge [2327]. Moreover, our new data augmentation and model aggregation methods [2324] deal directly with the problems of non-IID data distribution upon which it is studied. Altogether, I can say that my study could be viewed as an important contribution to the research on federated learning within the medical data processing setting.

Our contribution of providing a viable implementation and discussing some of the problems in the literature gives us the path to the further research in how to develop more specialized federated learning systems that will be able to overcome the difficulties regarding medical data, privacy protection, and the best performance of models.

5.3 Implications for the Field of Biomedical Research

In this section, the findings of the research are explained and their relevance to the area of biomedical research are explained. The concept of applying the federated machine learning methodology to the research of diabetic foot ulcers and semantic segmentation of RGB images is addressed. It explores the privacy-sensitive nature of federated learning that can provide the opportunity to cooperate with healthcare facilities and create valid and powerful models without infringing on the privacy of patients [1-5].

The tendencies of future research, the development of the multidisciplinary collaboration and the extrapolation of the findings of the study regarding the question of the application to the sphere of biomedical engineering are also discussed. Over the last few years, research into the application of machine learning, specifically, deep learning, to the analysis of medical pictures has grown by a significant margin. The literature has largely been based on the centralized training paradigm where all the information is collected and processed in a single location. The structure, which most researchers are commonly interested in, is the convolutional neural networks (CNNs) and the methods have demonstrated an encouraging outcome concerning the accuracy and effectiveness.

We are not the only research that is in the process of attempting to employ federated learning to semantically segment diabetic foot ulcers. Federated learning method is more decentralized and privacy-preserving in nature, and therefore it is especially appropriate in the case of medical problems where the sensitivity of data is a major concern [7]. This will be an important contribution to the sphere of federated learning since previous studies have focused primarily on simpler issues and smaller datasets. Our model of federated learning performance measures are very similar to the ones published in earlier centralized models. It is especially fascinating as it proves that despite the issues associated with decentralized data and the possible high cost of data connectivity, the federated learning method does not impair the accuracy [32].

5.4 Limitations of This Study

The truth is that in any study there are weak points and so does this research. In this case, I would like to present the key issues which may have influenced the findings, so that nothing will be hidden. We are discussing the dataset we worked on, the techniques we employed, how we constructed the model and any shortcuts or assumptions that we made during the process. Once we confront them directly, then it becomes simpler to determine what must be fixed next and where we can continue working. The application of federated learning to semantic segmentation of diabetic foot ulcers is certainly a good prospect, but the boundaries should not be overshadowed [4]. To begin with, the transfer of data between devices or institutions is associated with a considerable headache: data heterogeneity. Although we took much time to prepare and augment the data, the spread and diversity of the dataset was likely to have influence on the effectiveness of the model and the extent to which we can rely on the data to make generalizations.

Add to that the fact that sites apply various imaging devices, lighting, or have different patient characteristics and you receive even greater inconsistency. That may interfere with the precision of segmentation. Since federated learning operates on the principle of maintaining a decentralized model, there exists excessive amounts of back and forth communication between the central (global) model and all the local ones. When you are working in an area where bandwidth is minimal or the internet connection is not very stable, that is a serious issue. We simulated federated learning in our case, in fact we did not have time to explore the implications of communication overhead using the HTTP protocol [9].

Also, all of our data was obtained in a single source. That is useful, but narrows down the big picture. Federated learning has a lot of applications in medical image processing, but we only have yet to understand its performance in scale to more diverse data. Another one- we may not be able to apply our results to a situation where you suddenly increase the number of participants or data distribution is highly variable. Our segmentation model is quite complex, which makes it pick up tricky patterns, but there is the cost of using complexity which causes its own headaches. These models are more computationally intensive and can over-fit particularly when the data is not diverse enough and is not necessarily representative of the larger population [32].

It is also not easy to match the resources available in the universities where it is taking part with the requirements of our model. And although federated learning is designed in such a way that it does not compromise privacy, complete privacy remains a challenging nutshell. It is never certain that whatever updates the model, there is always a possibility that it may leak the information, so-called inference attacks, so that is something that we should keep a close eye on.

CHAPTER 6

CONCLUSSION AND FUTURE WORK

6.1 Conclusion

This study adopted the deep learning method to identify diabetic foot ulcers in the medical images- and it was successful. The team demonstrated that the automated analysis of images through the application of convolutional neural networks and transfer learning has a significant effect on the accuracy and reliability of DFU detection.

Their focus was to make the most out of a VGG16 based model so that it could give good diagnostic outcomes without dragging things down with heavy duty computing. It implies that their approach can indeed be applied in the biomedical context of reality. They did not simply stop at creating a model. They make much effort in preprocessing and data augmentation to clean the pictures and labor about the edges of their dataset. That was a significant difference that helped the model to be more generalized and the model was not overfitted. The model was found to perform well across the board on all the measures when they tested it, accuracy, precision, recall, F1-score. In essence, it was able to distinguish with minimal difficulty between ulcerated and healthy feet [8]. Even in occasions where they ran the confusion matrices down to the ground, misclassification hardly ever arose. An aspect that stands out the most about this study is the fact that they added explainable AI tools such as Grad-CAM and LIME.

Using these they were able to literally demonstrate the areas of the images that the model was paying attention to and can underline the areas that provide important information to the clinicians. This corresponds well to those visual explanations and the actual ulcer locations that comes across as trustworthy and increases the transparency of the system. Such confidence is immense in case AI is to enter the realm of clinical practice. This study demonstrates that deep learning models indeed can be used to identify diabetic foot ulcers early. Certainly, the road might have some additional bumps yet, such as dissimilarities in datasets and image quality, the system develops a strong foundation to advance in the future [5].

Starting it on larger and more diverse datasets takes the matter to the extreme. The method is already presentable as an actual application, be it in the clinics, mobile health applications, or telemedical [1]. Bottom line: this research will take the needle forward towards more intelligent, transparent, and effective methods of identifying diabetic foot ulcers in the current healthcare.

6.2 Summary of Findings

This paper investigated the application of deep learning to identifying diabetic foot ulcers on the medical images. We primarily aimed at creating a system that would be not just precise and rapid but easily interpreted by doctors which would in fact allow in the early diagnosis and decision making. The outcome is self-explanatory: convolutional neural networks, particularly when we had VGG16-based transfer learning model, were quite effective in identifying diabetic foot ulcers in the pictures.

A close data preparation and clever enrichment turned out to be significant to us. Mundane actions such as resizing, normalization, noise reduction followed by a shake up with rotation, flips and brightness adjustments actually assisted the model in learning what was important and prevented the model to get stuck with peculiarities of the training data [5]. That was important since medical image datasets are often not very big and unbalanced that tends to be problematic. When we put the model to the test, the numbers looked good throughout. Good accuracy indicated that it could correctly classify most images and the precision and recall scores indicated that it was reliably able to detect ulcers without having too many false alarms.

This was supported by the F1 score which depicted a reasonable tradeoff between catching ulcers and over-predicting. The confusion matrices narrated the same story- majority of images were found under the right category. Another point was another, namely VGG16 model, even when reduced and only further refined, could still make a stand when compared to other heavier models [21]. What that implies is that it is feasible in environments where computing capacity is low such as in a mobile phone or even in remote clinics. It is not high-tech as an end in itself, actually, this method can be applied to the real world.

6.3 Recommendations for Future Research

This paper demonstrates that deep learning is effective in the detection of diabetic foot ulcers, but it is not the end yet. Next steps? Begin with larger and more diversified data sets in other hospitals and regions. That will make the models harder and stronger [15]. It also assists in making use of the images that are captured in various lighting conditions, on various cameras and in varying resolutions- in short, as close as possible to what physicians see in the clinics. More recent deep learning systems are worth delving into [33].

Hybrids or transformer-based models that incorporate CNNs can assist the system to pick up finer details of the ulcer images. And do not limit yourself to yes/no answers, see how multi-class classification will help classify ulcer stages, level of infection, or vascular problems. That provides physicians with more valuable information. The addition of other data of various types, such as patient age, medical history, or even sensor data, enhances the prediction of the model and drives things towards more individualized care. To get these tools to operate in real time, these models must be able to operate on phones or edge devices, and not only on large servers. And the final thing is that people have to be confident in these systems [17].

It implies simplifying the decisions made by the AI and checking all that on actual healthcare specialists. AI in diabetic foot ulcer detection with its transparent and open explanations and good clinical testing can literally become a part of medical practice and be safe and transparent.