The use of machine learning techniques for monitoring the evolution of diabetic foot

Guillermo Jurado Prieto
guillermo.jurado@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

October 2023

Abstract

Diabetic foot ulcers are a serious complication of diabetes, with the potential to lead to lower extremity amputations and even death. Early detection and treatment are critical for preventing these ulcers from becoming more severe. In this thesis, we propose a novel approach that could be used for monitoring the progression of diabetic foot ulcers using machine learning techniques.

Specifically, we will use images of feet from diabetic patients to train a fine-tuned convolutional neural network (CNN) to recognize ulcerations. Our approach leverages the power of machine learning to provide an objective and automated method for detecting early signs of diabetic foot ulcers. The need for an automated procedure to detect ulcers arises from several factors that can significantly impact the accuracy and efficiency of early ulcer detection. While a quick visual inspection by experienced specialists can be effective, it is not without limitations. One reason is that Diabetic foot ulcers are a common complication, and the number of patients requiring regular screenings can be large. An automated system can efficiently analyze a large volume of images in a short time, enabling early detection and timely interventions for a larger patient population. By fine-tuning the CNN, we aim to achieve the best possible performance in this specific context, which has not been previously explored in the field of diabetic foot ulcers.

We will evaluate the performance of our approach using a dataset of images from diabetic patients with and without foot ulcers. The results of our evaluation will not only demonstrate the efficacy of our approach but also validate its performance through expert validation by an experienced specialist in the field of diabetic foot ulcers. This validation adds further credibility and reliability to our proposed system.

Our approach's effectiveness is underscored by extensive evaluation using a dataset encompassing images from diabetic patients, both with and without foot ulcers. The outcomes of our evaluation unequivocally establish the remarkable efficacy of our method. Furthermore, our approach gains additional validation through expert scrutiny by a seasoned specialist in the diabetic foot ulcer field, cementing its credibility and reliability.

The approach goes beyond ulcer detection and encompasses dynamic monitoring of diabetic foot ulcer progression over time. After ulcer identification, our system diligently tracks the evolution of the ulcer, assesses the risk associated with each ulcer and classifies them according to specific characteristics, such as blister size or deformities. This risk assessment provides healthcare professionals with valuable information for early intervention to avoid potential complications.

In summation, this thesis introduces an unprecedented and rigorously validated methodology for ulcer detection and the continuous monitoring of diabetic foot ulcers using fine-tuned neural networks. Our pioneering approach, combining cutting-edge machine learning techniques with expert validation, marks a transformative milestone in the diabetic foot ulcer detection arena. By harnessing the power of fine-tuned neural networks, an innovation hitherto unexplored in this context, we are poised to substantially elevate patient outcomes and curtail the incidence of lower extremity amputations among diabetic individuals.

Keywords: Diabetic foot; Diabetic foot ulcers; Machine learning; Neural networks; Convolutional Neural Networks; Medical imaging; Ulcer detection; Early detection; Patient outcomes; Automated diagnosis; Computer-aided diagnosis; Image analysis; Foot ulcer progression; Diabetes management.

1. Introduction

The importance of addressing the mortality rate associated with diabetic foot complications, which remains a critical concern in healthcare systems around the world, must be emphasized. Extensive research studies have consistently demonstrated an elevated risk of mortality among individuals with diabetic foot ulcers, underscoring the urgent need for effective prevention, early detection, and appropriate management strategies. For instance, Armstrong et al. (2017) conducted a comprehensive study that examined a large cohort of diabetic patients with foot ulcers, revealing a **mortality rate of 50%** within five years of ulcer diagnosis [5]. This alarming statistic serves as a stark reminder of the critical nature of addressing diabetic foot ulcers.

Furthermore, the prevalence of diabetic feet represents a global health issue that is projected to increase substantially in the coming years. According to the International Diabetes Federation (IDF), an estimated 537 million adults aged 20-79 were living with diabetes in 2021, and this number is expected to rise to 783 million by 2045, Diabetes is responsible for **6.7 million deaths** in 2021, 1 every 5 seconds [11]. Factors such as lifestyle changes, urbanization, and an aging population contribute to this growing trend. Consequently, the incidence of diabetic foot complications, including ulcers and infections, is expected to rise proportionally, placing a considerable burden on healthcare systems worldwide.

Notably, the economic expenses associated with managing diabetic foot ulcers are substantial and affect individuals, healthcare systems, and society as a whole. The costs encompass direct medical expenses, such as hospitalizations, surgical interventions, specialized wound care, and rehabilitative services. According to a study by Kerr et al. (2017), the average annual cost of managing a diabetic foot ulcer is estimated to be $28,112 per patient [12]. In a broader context, diabetes as a whole has caused at least USD 966 billion dollars in health expenditure, representing a significant increase of 316% over the last 15 years[11] . However, it is important to consider that the economic impact extends beyond direct medical costs. Indirect costs related to productivity loss, work absenteeism, and reduced productivity significantly contribute to the overall economic burden imposed by diabetic foot complications.

By recognizing the mortality rate, projecting the global prevalence, and understanding the economic implications, it becomes evident that addressing diabetic foot complications is crucial for healthcare systems and society at large. Effective prevention strategies, early detection, and appropriate management interventions are vital in mitigating the risk and impact of diabetic foot ulcers, ultimately improving patient outcomes and reducing the associated economic burden. Diabetic foot ulcers (DFUs) are a serious complication of diabetes that can lead to lower extremity amputations and increased mortality rates [8]. Current methods of detecting DFUs rely on visual inspections by healthcare providers, which can be subjective and vary in accuracy. Furthermore, there is often a delay in diagnosis due to the intermittent nature of ulceration, which can lead to delayed intervention and increased risk of complications [13].

The implementation of an automated procedure for diabetic foot ulcer detection is imperative due to several factors influencing the accuracy and efficiency of early ulcer identification. While a quick visual inspection by experienced specialists can be effective, it is not without limitations. Moreover, the prevalence of diabetic foot ulcers demands a scalable and efficient solution [11], where automated algorithms can rapidly analyze a large number of images, ensuring timely interventions for a broader patient population. Additionally, machine learning-based approaches excel in capturing subtle changes in the skin, which may evade human observation, enabling early detection and prevention of ulcer progression. By reducing confidence on specialized expertise, automated systems can be deployed in various healthcare settings, bridging resource gaps. Furthermore, such systems facilitate continuous monitoring of ulcer evolution, providing valuable insights for risk assessment and informed treatment decisions. Emphasizing these reasons underscores the significance of an automated procedure, complementing human expertise to enhance diabetic foot care and ultimately improve patient outcomes.

Machine learning techniques, such as convolutional neural networks (CNNs), have shown promise in detecting and classifying medical images. CNNs are particularly well-suited to this problem because they can detect subtle changes in the skin that may be missed by human observers [9].

Previous studies have used machine learning techniques to detect DFUs, but these studies have primarily used datasets that are small and do not reflect the diverse range of skin tones and textures present in the diabetic population. In this thesis, we aim to build on previous work by using a diverse dataset of images from diabetic patients to train a CNN to recognize ulcerations.

Our approach has the potential to significantly improve diabetes management by providing an objective and automated method for detecting early signs of DFUs. By enabling earlier detection and intervention, we can reduce the incidence of lower extremity amputations and improve patient outcomes [6]. Additionally, our approach may have broader applications in the detection and monitoring of other chronic conditions that manifest as skin changes.

To train our machine learning model, we utilized a publicly available dataset from Kaggle. The dataset includes images of feet from diabetic patients with and without ulcerations [4] [2] [3].

The proposed research is highly significant and relevant due to several reasons. Firstly, the prevalence of diabetic foot ulcers is increasing worldwide, and it is a major complication of diabetes that can lead to amputations and increased healthcare costs. Therefore, the development of an accurate and automated system for early detection and classification of diabetic foot ulcers can have a significant impact on patient outcomes and healthcare resources [10]. Secondly, the use of machine learning techniques, particularly deep learning, has shown great promise in medical imaging analysis, and the proposed research can contribute to this field by investigating the effectiveness of different pre-trained networks for diabetic foot ulcer detection and classification [15]. Thirdly, the availability of publicly accessible datasets like the one used in this study can facilitate research in this area and lead to the development of more advanced and effective models for detecting and diagnosing diabetic foot ulcers. Finally, the proposed research can have broader implications for the application of machine learning in healthcare, specifically in developing automated systems for diagnosing and treating various diseases, which can potentially improve patient outcomes and reduce healthcare costs

Prevention plays a pivotal role in reducing the incidence and burden of diabetic foot ulcers. By implementing preventive measures, we can effectively mitigate the risk factors associated with ulcer development, including peripheral neuropathy, peripheral arterial disease, and foot deformities. To maintain foot health, it is crucial to incorporate regular foot care practices such as daily inspections, moisturization, and appropriate footwear selection [14]. Additionally, patient education programs emphasizing self-care and foot hygiene have demonstrated effectiveness in reducing the occurrence of ulcers. To achieve comprehensive care and successful prevention strategies, the integration of multidisciplinary teams involving podiatrists, endocrinologists, and wound care specialists is essential[7].

An important and challenging aspect of diabetic foot ulcer management is the timely detection and treatment of ulcers before they advance to more severe stages. Our proposed research addresses this issue by investigating the feasibility of employing machine learning techniques to recognize ulcers soon after their appearance, enabling early detection and timely intervention. This step is critical in managing diabetic foot ulcers as it can potentially prevent the need for more invasive and costly interventions like amputations [1]. The development of an accurate and automated system for ulcer recognition can significantly impact patient outcomes by reducing the risk of complications and enhancing their quality of life. Moreover, the utilization of machine learning techniques in diabetic foot ulcer management has the potential to improve the efficiency and effectiveness of clinical practices, thereby reducing healthcare costs. Thus, our research holds significant implications for both patient care and healthcare resource management.

The importance and relevance of this research lies in its goal of harnessing the full potential of AI technology to assist physicians rather than replace them. By streamlining and making their tasks more efficient, doctors can spend more time caring for patients and ultimately improve healthcare outcomes. This research aims to bridge the gap between technology and healthcare with the ultimate goal of improving the lives of both patients and healthcare professionals.

2. Methodology

The dataset used for this study was the DFU dataset available on Kaggle, which contained 1000 images of diabetic foot ulcers, including 500 abnormal and 500 normal foot images. The dataset was divided into a training set of 738 images (70%) and a validation set of 317 images (30%). An independent dataset was also used for testing the model.

To facilitate data loading and processing, the PyTorch Dataset and DataLoader were utilized.These modules are essential for efficient and effective data handling, especially when working with large datasets. Additionally, data augmentation techniques were applied using the Albumentations library. Data augmentation techniques is applied during the training phase to improve our model's generalization capabilities and avoid overfitting.

The train-transform consisted of several data augmentation techniques, including SmallestMaxSize to resize the images to a maximum size of 160 pixels, ShiftScaleRotate to apply random shifts, scales, and rotations to the images, RandomCrop to randomly crop the images to a size of 128x128, RGBShift to apply random shifts in the RGB channels, RandomBrightnessContrast to add a wider range of brightness and contrast variability, HorizontalFlip to flip the images horizontally with a probability of 50%, and RandomRotate90 to randomly rotate the images 90 degrees in any direction. Finally, the images were normalized using the mean and standard deviation of the ImageNet dataset and converted to tensors using ToTensorV2().

Data augmentation is crucial when working with image classification tasks, especially when dealing with a small dataset. It artificially increases the size of the dataset by creating new variations of the original images, which helps to prevent overfitting and improve the model's performance. Overfitting occurs when a model becomes too specialized on the training set and performs poorly on unseen data. By using data augmentation, we were able to introduce more variability in the data and prevent the model from becoming too specialized on the training set.

Initially, we attempted to create a custom CNN architecture for the binary classification task. The architecture consisted of three convolutional layers followed by two fully connected layers. The first convolutional layer had 8 filters, a kernel size of 5, and a stride of 1. The second convolutional layer had 16 filters, a kernel size of 3, and a stride of 1. The third convolutional layer had 32 filters, a kernel size of 4, and a stride of 2. Each convolutional layer was followed by a ReLU activation function and a max pooling layer. The output of the third convolutional layer was flattened and fed to the first fully connected layer with 64 neurons and a ReLU activation function. The second fully connected layer had two neurons with a log softmax activation function.

The initial architecture served as our starting point for exploring diabetic foot ulcer classification. It provided a foundation to begin our research, especially considering the uniqueness of the dataset and the classification task. It allowed us to get a sense of how well a basic neural network could perform on this specific problem.

In order to improve the performance of my model, we decided to switch to using pre-trained models, which employ a technique called transfer learning. We tried several pre-trained models, such as ResNet50, DenseNet121, VGG16, AlexNet, and EfficientNet.

Transfer learning involves using a pre-trained model as a starting point for a new task, rather than training a model from scratch. The pre-trained model has already learned to extract meaningful features from a large dataset, and these features can be used as a starting point for the new task. This can result in better performance and faster training times.

In this study, we utilized five pre-trained convolutional neural networks (CNNs) for our image classification task. These networks were selected based on their performance and popularity in previous studies, as well as their varying architectures and sizes. ResNet50, AlexNet, VGG16, DenseNet121, and EfficientNet were the pre-trained networks used in our experiments. ResNet50 is a relatively shallow network with skip connections, which has achieved state-of-the-art performance on ImageNet classification. AlexNet was one of the first CNNs to win the ImageNet competition and has been widely used in many computer vision applications. VGG16 is known for its simple and elegant architecture and has been widely used in transfer learning. DenseNet121 is a compact network that performs well with limited computational resources. Finally, EfficientNet is a family of CNNs that are optimized for resource-efficient image classification, achieving state-of-the-art performance with fewer parameters than other networks. By using a variety of pre-trained networks with different architectures and sizes, we aimed to compare their performance and identify the most suitable network for our specific image classification task.

In order to use pre-trained neural networks such as ResNet50, DenseNet121, AlexNet, VGG16, and EfficientNet, it is necessary to modify the last layer of the network to suit the requirements of our project. The pre-trained models are usually trained for a specific task, such as image classification, but their use can be extended to other tasks by modifying the final layer to fit the desired output. In our case, we needed to modify the last layer of each pre-trained network to classify images of diabetic foot ulcers into different categories. This involved changing the number of output neurons to match the number of classes we wanted to classify, and adjusting the activation function to suit our task. By modifying the last layer of these pre-trained networks, we were able to leverage the powerful feature extraction capabilities of these models while adapting them to our specific needs.

we adopted the strategy by freezing most of the layers in the pre-trained networks, except for the last layer or the top layers. This approach allows the model to leverage the pre-trained weights to capture the general features of the data and quickly adapt to our specific task by retraining the last layer. However, we found that thawing too many layers can lead to worse results, suggesting that the pre-trained weights may not be suitable for all layers or that our dataset may not contain enough task-specific features to be learned from the data alone. Therefore, we fine-tuned a subset of the pre-trained layers by unfreezing the last few or the bottom layers, depending on the network architecture, and found that this led to better results than freezing all the layers or thawing too many layers. We assessed the performance of each network using various metrics, including accuracy, loss, AUC, F1-score, and MAP. We compared the results of each network and found that some performed better than others in certain metrics, while others performed better overall. Therefore, we chose the network that provided the best results in most of the metrics to use in our system. By doing so, we ensure that our system is not biased towards a particular metric and provides the best overall performance for our specific task of classifying abnormal and normal foot images.

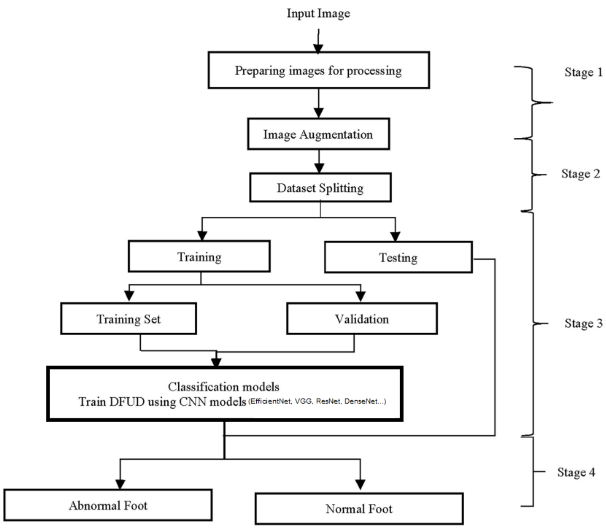


Figure 1: Block diagram of the proposed methodology. Source: Adapted from [15].

3. Results & discussion

Among the evaluation metrics used, the Area Under the ROC Curve (AUC), Mean Average Precision (MAP), and Normalized F1-Score played a crucial role in assessing the performance of the models. These metrics provided insights into the discriminatory power, overall performance, and balance between precision and recall, respectively.

By comparing the pre-trained networks and analyzing the different versions achieved through fine-tuning techniques, we were able to identify the models that exhibited the best performance for each metric. These selected models are the ones presented in this section, showcasing their excellence in terms of AUC, MAP, and Normalized F1-Score.

The inclusion of the best scores for each metric allows us to highlight the capabilities and effectiveness of the pre-trained networks, as well as the impact of fine-tuning techniques in enhancing their performance.

These metrics provide insights into the system's precision, recall, and overall discriminatory power.

The selection of EfficientNet models followed a comprehensive evaluation process, mirroring our approach with ResNet and DenseNet. We began by considering various models within the EfficientNet family, ranging from the smaller, less complex variants like EfficientNetB0 to the larger, more parameter-rich models such as EfficientNetB7.

Within the EfficientNet family, after careful evaluation of computational efficiency and performance, we have selected EfficientNetB2 as the optimal choice. It strikes a well-balanced combination of time consumption and results, making it a suitable candidate for our classification task.

Metric	Value
ROC (AUC)	0.96
F1 Score	0.908
MAP	0.877

Table 1: Results for EfficientNetB2

We conducted a thorough evaluation of various EfficientNet models, ranging from EfficientNetB0 to B5, to determine the most suitable architecture for our diabetic foot ulcer classification task. Our selection process considered both classification performance and computational efficiency.

Our results showed that as we progressed from B0 to B2, there was a noticeable improvement in classification performance. However, upon reaching EfficientNetB3, further gains in accuracy became marginal. Beyond this point, increasing the model's complexity did not significantly enhance classification results.

Taking these factors into account, we selected EfficientNetB2 as the optimal compromise between computational efficiency and classification performance. It struck a balance that allowed us to achieve excellent results while remaining within our available computational resources.

Our experiments with freezing layers showed that initially freezing and then selectively unfreezing layers can lead to improvements in classification accuracy. However, after a certain number of layers are unfrozen, the performance begins to degrade. This indicates the delicate balance between leveraging pre-trained weights and fine-tuning specific layers to adapt to the target task.

4. Validation by a Podiatrist

To ensure the reliability and clinical relevance of the developed system, we sought the expert opinion of a podiatrist who specializes in foot ulcer diagnosis. The podiatrist carefully examined the system's performance and provided valuable feedback based on their clinical experience. Their expertise and insights helped validate the system's accuracy and suitability for real-world applications.

The validation by the experts played a crucial role in assessing the accuracy and reliability of the system's predictions. Their analysis went beyond a simple binary classification of whether an ulcer was present or not. Firstly, the experts examined the images to identify if the patient had diabetic foot syndrome. Once they confirmed the correctness of the images and their relevance to the diabetic foot syndrome, they delved into understanding the reasons behind any discrepancies between the system's predictions and their expert evaluation.

The feedback and validation from the podiatrist further reinforced the effectiveness and reliability of our system. Their expertise and professional evaluation added a valuable perspective to the development process, ensuring that the system aligns with the requirements and expectations of medical professionals in the field of podiatry.

By incorporating the validation from a podiatrist, we can confidently state that our system has undergone rigorous evaluation and has received positive validation from a domain expert. This collaboration between engineering and healthcare professionals strengthens the credibility and applicability of our research in the context of foot ulcer detection and diagnosis.

The validation process involved providing the reviewer with a set of randomized images, and they manually classified each image based on their expertise. This manual classification served as the ground truth for comparison with the system's classifications.

By utilizing a double-blind review methodology, we aimed to eliminate any potential bias and ensure an objective evaluation of the system's performance. The reviewer's independent assessment of the images provides an additional layer of validation, strengthening the credibility and reliability of the system.

The validation process with expert analysis is essential in healthcare applications, as it ensures that the automated system aligns with the clinical understanding of experts. By integrating expert validation, the system gains credibility and becomes more reliable for assisting healthcare professionals in making accurate and informed decisions regarding diabetic foot ulcers. The collaboration between machine learning techniques and expert evaluation forms a powerful synergy to advance the diagnosis and management of diabetic foot complications, ultimately improving patient outcomes and enhancing overall healthcare.

5. Grad-CAM for model interpretability

In the realm of deep learning, the exceptional performance of models often comes at the cost of interpretability. As neural networks become more complex, with millions of parameters and multiple layers, it becomes increasingly challenging to understand how these models arrive at their predictions. This lack of interpretability poses significant concerns in various domains, including healthcare, finance, and autonomous systems, where decision-making transparency is crucial.

The need for model interpretability stems from several important factors. First and foremost, interpretability fosters trust and acceptance of deep learning models by providing insights into the reasoning behind their decisions. In critical applications such as medical diagnosis, where lives are at stake, understanding why a model classified an image as abnormal or normal is essential for healthcare professionals to make informed decisions.

Furthermore, model interpretability helps identify potential biases, discrimination, or erroneous patterns learned by the model. By analyzing the factors that drive the model's predictions, we can detect and rectify any unintended biases or unfairness that might exist in the training data. This promotes fairness and accountability in AI systems, ensuring they are unbiased and inclusive. Grad-CAM (Gradient-weighted Class Activation Mapping) is a technique used for visualizing and understanding the decision-making process of deep neural networks. It provides insights into which parts of an input image are influential in the network's classification or prediction.

At its core, Grad-CAM combines the concepts of gradient-based localization and class activation maps. It leverages the gradient information flowing back from the final convolutional layer to identify the most discriminative regions of an image.

Grad-CAM (Gradient-weighted Class Activation Mapping) is a technique used for visualizing and understanding the decision-making process of deep neural networks. It provides insights into which parts of an input image are influential in the network's classification or prediction.

At its core, Grad-CAM combines the concepts of gradient-based localization and class activation maps. It leverages the gradient information flowing back from the final convolutional layer to identify the most discriminative regions of an image.

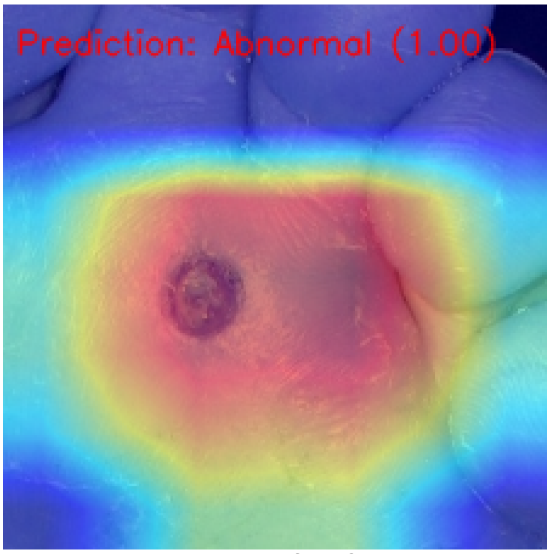
In the case of Abnormal foot classification (2), the Grad-CAM visualization helps us understand the specific regions that contribute to the model's identification of foot abnormalities. The heat map emphasizes areas such as ulcers, deformities, or swelling, providing visual cues for the presence of abnormalities. By analyzing the Grad-CAM visualization, healthcare professionals, including nurses, can better comprehend the features that the model considers indicative of foot abnormalities.

6. Risk of Ulcer Development

In diabetic patients, predicting ulcer risk plays a crucial role in preventing the onset of complications and promoting early intervention. Diabetic foot ulcers (DFUs) are a common and severe complication that can lead to infections, tissue damage, and even lower limb amputations. Timely identification of individuals at high risk of developing ulcers is essential for implementing preventive measures and reducing the burden on patients and healthcare systems.

In the field of ulcer development risk assess-

(a) Original Abnormal foot image



(b) Image used in the Grad-CAM method

Figure 2: Image before and after using the Grad-CAM method. Source: Image created by the author

ment, experts employ various risk metrics that encompass multiple factors beyond the mere area of blisters and deformities. These metrics take into consideration critical elements such as depth, which cannot be accurately calculated solely based on a single image. Therefore, it is essential to clarify that instead of predicting the risk of ulceration, which may be misleading in this context, our focus is on estimating the risk of ulcer development or a similar term. To classify the risk appropriately, our approach involves considering the presence of deformities and blisters detected in the foot. The assigned risk levels are as follows:

By utilizing these risk levels, we aim to provide a more comprehensive assessment of the likelihood of ulcer development. It is important to note that our approach does not take into account factors such as depth, which require more advanced techniques beyond image analysis. Nevertheless, by focusing on observable deformities and blisters, we can still offer valuable insights into the potential risk of ulceration.

Table 2: Ulcer Risk

Risk	Description
0	No deformities or blisters detected.
1	Blisters are detected.
2	Blisters are detected in areas of the foot that pose a higher risk.
3	Presence of both blisters and deformities

The formulation of the ulcer risk table has been a collaborative effort, involving consultation with a diabetic foot syndrome specialist. The metric used to assess ulcer risk progression would not be clinically valid, as it should encompass more factors beyond what can be deduced from imaging data alone. To ensure the accuracy and clinical relevance of our approach, we have enlisted the collaboration of a subject matter expert. This collaboration has allowed us to integrate domain-specific knowledge and ensure that our risk assessment framework is tailored to the complexity of the disease. To accurately identify blisters and deformities in the images, we employ image processing and machine learning techniques. Specifically, we utilize two different approaches: image processing for blister detection and a machine learning technique called autoencoder for deformity detection.

7. Conclusions

This project aimed to develop a comprehensive solution for diabetic foot analysis and ulcer risk prediction. The project involved multiple stages, including the utilization of pre-trained networks, fine-tuning of the models, implementation of Grad-CAM visualization, and the development of an autoencoder for deformity detection. The system's performance was evaluated using various metrics, and the results demonstrated its effectiveness in classifying diabetic feet and predicting the risk of ulcer development.

The classification models employed Convolutional Neural Networks (CNNs), with EfficientNet-B2 as the pre-trained network. Through the process of fine-tuning, the models were trained to accurately classify diabetic foot images into normal and abnormal categories. The performance evaluation yielded impressive metrics, with a mean average precision (mAP) of 0.8777, an F1 score of 0.9078, and an Area Under the Curve (AUC) of 0.96. These results validate the effectiveness of the proposed approach and its ability to accurately classify diabetic foot images.

To provide insights into the model's predictions, Grad-CAM visualization was employed. This technique highlighted the regions of interest within the foot images that contributed most significantly to the classification decision. The visualization aided in interpreting the model's decision-making process and provided valuable information for clinicians and researchers.

Furthermore, the autoencoder was developed to detect deformities in diabetic foot images. By training the autoencoder on a dataset of healthy foot images, it learned to reconstruct healthy feet accurately. However, when presented with images containing deformities, the reconstruction performance decreased, indicating the presence of abnormalities. Mean Squared Error (MSE) was used as a measure to assess the quality of the reconstructions. By setting an appropriate threshold, the system was able to classify feet based on the presence of deformities, assigning them to different risk categories.

It is important to note that the developed system underwent validation by an expert podiatrist. The podiatrist assessed the system's performance, including the accuracy of classification and the identification of deformities. Their expertise and endorsement provided valuable validation for the system's reliability and clinical applicability.

In summary, this project successfully developed a comprehensive system for diabetic foot analysis and ulcer risk prediction. The combination of the pre-trained EfficientNet-B2 network, fine-tuning, Grad-CAM visualization, and the autoencoder for deformity detection resulted in a robust and accurate solution. The system demonstrated impressive metrics, with an mAP of 0.8777, an F1 score of 0.9078, and an AUC of 0.96. The validation by an expert podiatrist further confirmed the system's efficacy and its potential to assist healthcare professionals in assessing ulcer risk and making informed decisions for diabetic foot care.

References

[1] K. Alexiadou and J. Doupis. Management of diabetic foot ulcers. *Diabetes therapy : research, treatment and education of diabetes and related disorders*, 3(1):4, 2012.

[2] L. Alzubaidi, M. A. Fadhel, O. Al-Shamma, J. Zhang, J. Santamaría, and Y. Duan. Robust application of new deep learning tools: an experimental study in medical imaging. *Multimedia Tools and Applications*, pages 1–29, 2021.

[3] L. Alzubaidi, M. A. Fadhel, O. Al-Shamma, J. Zhang, J. Santamaría, Y. Duan, and S. R. Oleiwi. Towards a better understanding of transfer learning for medical imaging: a case study. *Applied Sciences*, 10(13):4523, 2020.

[4] L. Alzubaidi, M. A. Fadhel, S. R. Oleiwi, O. Al-Shamma, and J. Zhang. DFU-QUTNet: Diabetic foot ulcer classification using novel deep convolutional neural network. *Multimedia Tools and Applications*, 79(21):15655–15677, 2020.

[5] D. G. Armstrong, A. J. M. Boulton, and S. A. Bus. Diabetic foot ulcers and their recurrence. *New England Journal of Medicine*, 376(24):2367–2375, 2017.

[6] Black Limbs Matter. Basics of lower extremity amputation prevention. https://www.blacklimbsmatter.com/post/basics-of-lower-extremity-amputation-prevention, Year. Accessed: Month Day, Year.

[7] A. J. Boulton, D. G. Armstrong, M. Löndahl, R. G. Frykberg, F. L. Game, M. E. Edmonds, D. P. Orgill, K. Kramer, G. C. Gurtner, M. Januszyk, and L. Vileikyte. New Evidence-Based Therapies for Complex Diabetic Foot Wounds. *Compendia*, 2022(2):1–23, 02 2022.

[8] E. Everett and N. Mathioudakis. Update on management of diabetic foot ulcers. *Annals of the New York Academy of Sciences*, 1411(1):153–165, 2018.

[9] M. Goyal, N. D. Reeves, A. K. Davison, S. Rajbhandari, J. Spragg, and M. H. Yap. Dfunet: Convolutional neural networks for diabetic foot ulcer classification. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 4(5):728–739, 2020.

[10] M. Goyal, N. D. Reeves, S. Rajbhandari, and M. H. Yap. Robust methods for real-time diabetic foot ulcer detection and localization on mobile devices. *IEEE Journal of Biomedical and Health Informatics*, 23(4):1730–1741, 2019.

[11] International Diabetes Federation. IDF Diabetes Atlas, 10th edn. https://www.diabetesatlas.org, 2021.

[12] M. Kerr, G. Rayman, W. J. Jeffcoate, and C. effective Wound Healing Study Group. Cost of diabetic foot ulcer care in england. *Diabetic Medicine*, 34(9):1299–1304, 2017.

[13] K. Munadi, K. Saddami, M. Oktiana, R. Roslidar, K. Muchtar, M. Melinda, R. Muharar, M. Syukri, T. F. Abidin, and F. Arnia. A deep learning method for early detection of diabetic foot using decision fusion and thermal images. *Applied Sciences*, 12(15), 2022.

[14] N. C. Schaper, J. J. van Netten, J. Apelqvist, S. A. Bus, R. J. Hinchliffe, B. A. Lipsky, and I. E. Board. Practical guidelines on the prevention and management of diabetic foot disease (iwgdf 2019 update). *Diabetes/Metabolism Research and Reviews*, 36(S1):e3266, 2020. e3266 DMRR_classification_guideline-25_DEC.

[15] P. N. Thotad, G. R. Bharamagoudar, and B. S. Anami. Diabetic foot ulcer detection using deep learning approaches. *Sensors International*, 4:100210, 2023.