GAN-Powered Data Augmentation for Robust Deep Learning Models In DFU Diagnosis

Tejal Khade
Computer Science Engineering
Sardar Vallabhbhai National Institute Of Technology
Surat, Gujarat, India
p23cs019@coed.svnit.ac.in

2nd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

3rd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

4th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

5th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

6th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

***Abstract*—Diabetic foot ulcers (DFUs) are a major complication of diabetes, often leading to severe consequences such as amputations if not diagnosed and managed promptly. Early and accurate classification of DFUs is essential for effective treatment and prevention of complications. Recently, a number of GAN variants have been emerged for improved quality in data augmentation. Although successful, further improvement is necessary for enhancing diversity in addition to quality in data augmentation.The advantage of GANs is to perform semi-supervised learning and unsupervised learning. In this paper, we propose a GAN-based approach to self-supervised augmentation of quality data based on Convolutional Neural Network(CNN) GAN referred to here as CNN GAN, to extending diversity as well as quality in data augmentation. This presents an innovative and comprehensive framework for the precise classification of diabetic foot ulcers using a deep learning-based architecture with generative adversarial networks. Our approach leverages advanced neural networks to automatically extract meaningful features from clinical DFU images, addressing the complex patterns and variations associated with different stages and types of ulcers. The proposed model integrates convolutional layers for hierarchical feature extraction, followed by fully connected layers to enhance feature representation. The framework is evaluated on a large and diverse dataset of DFU images, achieving superior performance compared to existing methods in terms of accuracy, precision, recall, and F1 score. This model offers a cost-effective, scalable, and efficient solution for DFU classification problem.**

***Index Terms*—Diabetic foot ulcers, deep learning, image classification, convolutional neural networks, attention mechanisms.**

I. Introduction

GAN frameworks provide an effective means for self supervised data augmentation with the fake data generated by their generators. To date, several GAN variants have been proposed for the sake of maximization of generated fake data's quality and diversities. Many GAN variants incorporate such data attributes as clusters and classes as well as self-supervised data augmentation into the adversarial framework, often, with a classifier attached to the discriminator with shared weights. ClusterGAN and its derivatives exploit the clustering structure embedded in fake data to enhance the quality in data generation with additional training of the generator based on an attached encoder. ACGAN, InfoGAN and their derivatives exploit the class structure and data attributes,instead, by using a classifier attached to the discriminator with shared weights. They achieve quality improvement in data generation by imposing the generator to follow class and attribute configurations embedded in the real and fake data. Note that, by "quality," we mean that the generated data are as real as the real data, while, by "diversity," we mean that the generated data genuinely cover the distribution of real data without bias. The state-of-the-art approaches introduced above have contributed, on their own sake, to the advancement of GAN-based data generation in terms of data quality as well as diversity. However, they are also subject to certain performance limitations due to the issues associated with goal inconsistency and constraint in data augmentation. A necessity to break through the current performance limit exists by overcoming the above issues. Goal inconsistency takes place in GAN frameworks due to additional learning of classes, categories and attribute distributions with a classifier shared its weights with the discriminator, which disturbs the adversarial framework to converge to real data distribution. Nonetheless, goal inconsistency remains as an issue to overcome for further advancement in GAN-based data generation. Such constraint in data augmentation serves as a limiting factor for self-supervised augmentation of real and fake data to improve diversity in GAN-based data generation. Therefore, achieving both high quality and high diversity in GAN based data generation remains a challenge.

GAN powered data augmentation can be used in many use cases like Diabetic Foot Ulcers (DFUs) in medical applications. DFU are a significant and common complication of diabetes, often leading to severe outcomes such as lower

Identify applicable funding agency here. If none, delete this.

limb amputation if left untreated. The global prevalence of diabetes has risen drastically from 108 million people in 1980 to over 422 million in 2014, and it is expected to reach 600 million by 2035. DFUs affect 15% to 25% of diabetic patients, with inadequate care resulting in amputations for many. The development of DFUs is driven by factors like neuropathy, poor circulation, and reduced immune response. As diabetes-related complications continue to burden individuals, especially in developing countries, there is an urgent need for effective diagnostic and treatment approaches to prevent further complications and reduce the financial impact on patients and healthcare systems.

In this paper, we propose an approach to GAN-based data augmentation that offers both high quality and high diversity in data generation.The proposed GAN framework is referred to here as "CLS GAN." an attached classifier to discriminator that corrects the fake data qualified for self-training the generator as well as for self-augmenting the real data. This architecture will further integrate convolutional neural networks to accurately classify Diabetic foot ulcers. GAN-based augmentation in medical analysis not only increases the diversity of the training data but also preserves the realism and underlying distribution of the medical images.

II. Literature Survey

A detailed literature survey is performed to study various approaches proposed by different researchers.

In [1], a deep neural architecture is proposed for the classification of diabetic foot ulcers (DFUs), addressing the unique challenges posed by DFU imaging data. The architecture integrates convolutional neural networks (CNNs) with residual blocks and feature fusion layers to enhance feature extraction and representation. The methodology includes data augmentation techniques such as rotation, flipping, Gaussian noise, shearing, and translation to improve model robustness. The model achieves remarkable results, with an accuracy of 98.87%, precision of 99.01%, recall of 98.73%, F1-score of 98.86%, and an AUC-ROC of 98.13% on benchmark datasets, including DFU2020 and MICCAI DFUs. However, limitations include the need for more diverse datasets, hyperparameter fine-tuning, and evaluation across different clinical scenarios to improve real-world applicability.

In [2], various advanced models were explored for DFU detection, including EfficientDet, Cascade R-CNN integrated with DetNet, and Faster R-CNN with deformable convolution layers. The study utilized data augmentation techniques such as random rotation and shear transformations to enhance the robustness of the models. Benchmark datasets, including the MS-COCO dataset and the Diabetic Foot Ulcers Grand Challenge (DFUC2020) dataset with designated training, validation, and testing sets, were employed. Despite promising results, challenges remain, such as optimizing CNNs for remote monitoring applications and reducing false positives caused by difficulties in distinguishing ulcers from other skin conditions.

In [3], the study presents the design of four hybrid CNN models for DFU classification, with an empirical comparison of different architectures featuring varying numbers of branches. The models utilized feature aggregation techniques, including a Global Average Pooling (GAP) layer and fully connected layers with dropout for enhanced classification performance. The final output was determined through a Softmax layer. The dataset used consisted of 754 images of patients' feet, collected from the Nasiriyah Hospital's diabetic center in Iraq, with two categories: abnormal (DFU) and normal (healthy skin). However, the study faced limitations, such as the inability to improve performance by increasing network width, the small size of the dataset, and the model's current focus on only two classes (normal vs. abnormal). Future work includes the application of the model for transfer learning.

In [4], the novel CNN model DFU_QUTNet was proposed for diabetic foot ulcer (DFU) classification and compared with state-of-the-art CNN architectures such as GoogleNet, AlexNet, and VGG16 after fine-tuning. The DFU_QUTNet, when combined with an SVM classifier, achieved a higher F1-Score of 94.5% compared to the other networks. The features extracted by DFU_QUTNet were utilized to train both SVM and KNN classifiers, with the SVM classifier demonstrating the highest precision, recall, and F1-Score. The dataset used in this study contained 754 feet images from patients with DFU, with each image region of interest (ROI) resized to 224x224. However, the study faced several limitations, including the limited dataset size of only 754 images, the lack of generalization to other tasks such as classifying skin cancer or detecting DFU in diverse cases, and the absence of clinical validation or comparison with expert clinicians.

In [5], an ensemble CNN model was compared against handcrafted machine learning algorithms for diabetic foot ulcer (DFU) classification tasks. The study focused on binary classification tasks, specifically differentiating between Ischaemia vs. Non-Ischaemia and Infection vs. Non-Infection. The dataset consisted of 1459 DFU images collected from the Lancashire Teaching Hospitals, with images captured using various camera models, including Kodak DX4530, Nikon D3300, and Nikon COOLPIX P100. Natural data augmentation techniques were employed to enhance the dataset. The study highlighted challenges such as high visual intra-class dissimilarities and inter-class similarities, imbalanced dataset, and limited data quality. Furthermore, there is potential for improving model performance by optimizing hyperparameters for both traditional machine learning algorithms and CNN models. One key limitation noted was that infection in DFU images may not always present clear visual indicators, which complicates the classification task.

In [6], an ensemble of Convolutional Neural Networks (CNN) and Vision Transformers (ViT) was proposed for diabetic foot ulcer (DFU) classification. The study also integrated a Siamese Neural Network (SNN) with a k-Nearest Neighbors (kNN) classifier for enhanced performance. K-Fold cross-validation was applied to ensure the robustness of the model. Data augmentation techniques were used to address dataset imbalance, improving generalization. The DFUC2021 Challenge dataset was used, which included classes such as "None" (no visible DFU), "Infection" (presence of bacterial or fungal infection), "Ischemia" (reduced blood flow causing tissue damage), and "Both" (co-occurrence of infection and ischemia). The dataset faced class imbalance, and pre-applied data augmentation was used to mitigate this. Despite these advancements, the study noted that it could not explore ensemble modeling or more computationally intensive variations to further enhance results.

In [7], the study addresses the growing challenge of Diabetic Foot Ulcers (DFUs), particularly those with ischemia and infection, emphasizing the need for early detection. The authors proposed using pre-trained transformer models, fine-tuned on the DFUC-21 dataset, for multiclass DFU classification.A Multi-Model approach was proposed, where features from parallel-trained transformers were fused from the last layers, achieving a macro-average F1-Score of 0.569. Weighted cross-entropy optimization and pairwise feature fusion addressed class imbalance. The results highlight the potential of combining CNNs with transformer architectures for future improvements in DFU classification.

In [8], the study highlights the importance of high-quality, diverse training data for deep learning applications and addresses challenges in obtaining sufficient realistic data. It proposes CLS-R GAN, a Classification-Reinforced GAN, to enhance both the quality and diversity of augmented data through a novel self-training framework. In CLS-R GAN, an independent classifier guides the generator to self-train by classifying fake data and augmenting real data in an unsupervised manner. Experiments, including applications to liver ultrasound image augmentation, demonstrated the effectiveness of CLS-R GAN in improving data quality and diversity. Future work aims to extend its application to datasets like CIFAR10, STL10, and ImageNet, and refine system parameters and fake data selection criteria for further optimization.

In [9], study evaluates the performance of multiple deep learning models for diabetic foot ulcer (DFU) detection, using a dataset of 640x480 pixel images annotated with LabelImg and VGG Image Annotator. The models include EfficientDet, which leverages EfficientNet as a backbone and BiFPN for feature fusion, YOLOv5, known for real-time object detection with single-pass processing, and Faster R-CNN, featuring a three-stage architecture for robust region proposals and detection. Data augmentation techniques such as scaling, color adjustments, and mosaic augmentation were applied to enhance diversity. The models were assessed using F1-score and mean average precision (mAP), highlighting EfficientDet's scalability, YOLOv5's speed, and Faster R-CNN's precision in detecting DFUs.

In [10], various techniques for addressing the challenges of diabetic foot ulcer (DFU) classification using the DFUC2021 dataset were performed. Given the dataset's class imbalance, performance evaluation included per-class F1-Score, micro-average F1-Score, and macro-averages of Precision, Recall, F1-Score, and AUC. Pretrained models from ImageNet and data augmentation strategies were employed to enhance model performance. For example, ischaemia images underwent eight augmentation techniques, while infection and ischaemia classes were augmented with three techniques, significantly increasing data samples. Among tested models, DenseNet121 and EfficientNetB0 emerged as top performers. DenseNet121 achieved the highest macro-average AUC of 0.88, while EfficientNetB0 excelled in macro-average Precision, Recall, and F1-Score, particularly improving the infection F1-Score. Analysis with UMAP highlighted EfficientNetB0's ability to enhance intra-class clustering on testing data, although inter-class separation remained challenging. Despite advancements, accurately detecting infection and co-occurrence of ischaemia and infection continues to be difficult, especially in the "both" category.

In [11], Conventional Machine Learning (CML) methods and Convolutional Neural Networks (CNNs) for ulcer versus non-ulcer classification were used. A novel CNN architecture called DFUNet is proposed, which is fine tuned to process the input data more effectively and efficiently than other comparative state-of-the-art CNNs architecture. Using 10-fold cross-validation, DFUNet achieved an AUC score of 0.961. This outperformed both the machine learning and deep learning classifiers which they have tested. CNNs architecture of LeNet, AlexNet, and GoogLeNet have been used to develop a fully automatic method to classify the DFU skin against the normal skin whereas their proposed DFUNet demonstrated superior performance compared to GoogLeNet and AlexNet in terms of accuracy and sensitivity. Data Augmentation techniques used were rotation, flipping, contrast enhancement, using different color space, and random scaling but that did not contribute to improve the overall performance.

III. Proposed Methodology

This section consists of five parts: (i) our dataset with samples of diabetic foot ulcers from diverse patients (ii) labeling process of our dataset into normal skin and abnormal patched skin categories (iii) preprocessing of training patches through data augmentation using GANs to improve model accuracy and use generative adversarial networks (GANs) by exploring various GAN models to synthesize realistic patches for normal and abnormal skin and comparing their classification accuracy from both techniques (iv) fine-tuned CNN architectures of EfficientNetB0 model as the base architecture, pre-trained on ImageNet dataset (v) our proposed model, a novel CNN architecture integrated with GAN principles to enhance DFU classification. Our proposed model employs GAN-based data augmentation to address class imbalance and increase data diversity while using its CNN module to extract robust features for accurate classification, achieving improved performance compared to fine-tuned pre-trained models.

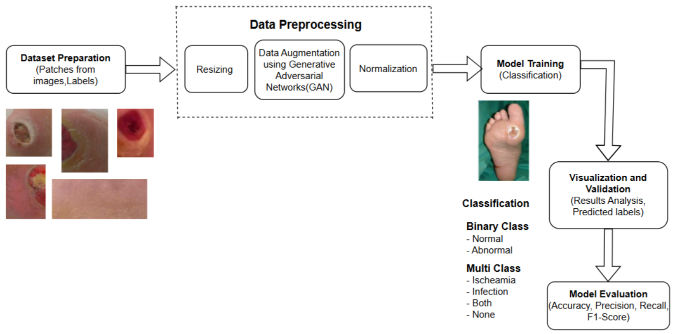

Fig. 1. Proposed Methodology

A. Dataset Preparation

We have received dataset (reference number is 332) to use diabetic foot ulcer (DFU) images for the purpose of this research. These images are photographs collected from the Lancashire Teaching Hospitals, where photographs were acquired from the patients during their clinical visits. Three cameras were used for capturing the foot images,Kodak DX4530, Nikon D3300 and Nikon COOLPIX P100. The images were acquired with close-ups of the full foot at a distance of around 30–40 cm with the parallel orientation to the plane of an ulcer. The use of flash as the primary light source was avoided, and instead, adequate room lights were used to get the consistent colours in the image. A podiatrist and a consultant physician with specialization in the diabetic foot helped to extract the images from the archive.

1) Ground Truth Annotation: The instruction for annotation is to identify the location of the ulcer with a bounding box and label each ulcer with ischaemia and/or infection, or none. They used annotation to identify the location of the ulcer with a bounding box and label each ulcer with ischaemia and/or infection, or none. They used the software, LabelImg, to label the images. Also, we first cropped the Region of Interest (ROI) with a size of 224×224. This region is a significant region around the ulcer which includes important tissues of both skin classes (normal and abnormal), then the specialist labeled the cropped patches. The ground-truth labels are marked by a medical specialist in two forms of normal and abnormal skin patches. We collected a total of 1679 skin patches with 641 normal and 1038 abnormal (DFU). Finally, we divided the data-set into 80% of patches for training, and 20% of patches for a testing set of images. There are multiple annotations from a podiatrist and a consultant physician with specialization in the diabetic foot. We average the bounding boxes to form a final bounding box. For diabetic foot pathology classification, the pathology labels were validated with medical records.

B. Data Pre-processing

The images in dataset are DFU patches of normal and abnormal skin which are of varying sizes and need to be converted into same size samples for training data. Thus, it is very important to perform preprocessing on these patches. The pre-processing pipeline begins by resizing the input images to 224x224 pixels, ensuring uniformity in dimensions for model input. Images are then center-cropped to retain the most relevant portions of the image. The pixel values are converted into tensors for efficient processing, with each pixel value normalized to fall within the range [-1,1]. This normalization helps improve model convergence during training. Finally, the preprocessed images are converted from tensor format back to NumPy arrays and transposed for visualization into (batch_size,height,width,channels) format.

C. Data Augmentation

A limited training dataset can lead to poorly tuned parameters and significant overfitting, reducing the model's generalization ability. The data augmentation process is a common approach to address this issue and has been shown to enhance the performance of various deep learning tasks. Additionally, collecting a large volume of medical data is often costly, time-consuming, and challenging due to privacy concerns and data scarcity. Many papers proposed traditional data augmentation techniques like Random rotations, Random translations, Random flips, Gaussian noise, Random contrast to generate images for classification. These techniques can be used when there is not much difference between the two classes and can be prevented using traditional augmentation techniques. Our dataset has high imbalance between two classes i.e Normal and Abnormal, thus we overcome this limitation by using Generative Adversarial Networks (GANs) to generate synthetic training patches, particularly for rare or underrepresented classes in the dataset which can prove to give better classification accuracy. The DFUC2021 dataset is imbalanced in terms of its class distribution. Thus we performed data augmentation using Generative Adversarial Network (GAN) to balance the dataset. We have discarded images with exact matched and minor scaling changes (generated by data augmentation using GAN). Average hashing and wavelets hashing are not suitable for sanity check, as they clustered images with inter-similarities. We ran a sanity check to reduce the repetition of similar images. We tested image hashing techniques [12], i.e. different hashing, average hashing, perceptual hashing and wavelets hashing to check the similarity of the DFU images. Different hashing finds the exact match of the images and perceptual hashing finds the images with different scales. We have discarded images with exact matched and minor scaling changes (generated by natural data augmentation).

D. Model Training and Data Distribution

For DFU classification, we followed a 80-20 split for training and testing set. We further split 20% from training set, which we used as a validation set. The total images for DFUC2021 is 15,683, with 5,955 images for training set (2,555 infection only, 227 ischaemia only, 621 both infection and ischaemia, and 2,552 without ischaemia and infection), 3,994 unlabeled images and 5,734 images for testing set.

E. Visualization and Validation

Visualization techniques play a crucial role in understanding model predictions, feature extraction, and dataset properties. Feature Map Visualizations can be used to understand which features the CNN is extracting from input images. Loss Curves from generator and discriminator can be plotted to track adversarial training progress. Generated Samples from the generator can be monitored regularly so as to check for improvements in realism and diversity.

F. Model Evaluation

The DFUC2021 dataset is imbalanced in terms of its class distribution. To properly handle such class imbalances, the performance is to be reported with per class F1-Score, micro average F1-Score and area under the Receiver Operating Characteristics Curve (AUC) and macro-average of Precision, Recall, F1-Score and AUC to reflect the overall performance. While the micro-average reflect the overall performance, the macro-average is a good choice in imbalanced multi-class settings [13] as it caters well for cases of strong class imbalance. Accuracy can also be determined for resultant binary classification.

G. Description of Workflow Design Architecture

An innovative architecture of a convolutional neural network(CNN) in combination with Generative Adversarial Network(GAN) is proposed to improve the extraction of critical features for diabetic foot ulcer (DFU) classification. Two significant issues are addressed in designing this architecture: (i) While increasing the number of convolutional layers in traditional CNN models can enhance accuracy up to a certain point, an excessive number of layers may lead to performance degradation due to overfitting or vanishing gradients. (ii) Simple networks with fewer layers may suffice for basic applications, but DFU classification demands a more intricate structure to capture the subtle differences between normal and abnormal classes effectively.

The proposed neural network, is a deep neural network architecture designed to efficiently process DFU images for classification tasks. It emphasizes deep convolutional layers, skip connections, bottleneck layers, and dense connections to optimize feature extraction while maintaining computational efficiency. Skip connections are skillfully integrated to allow gradients to travel efficiently across layers, mitigating the vanishing gradient problem and enhancing the network's ability to capture long-range dependencies. Dense connections further facilitate seamless information flow between layers, improving gradient propagation and feature reuse.

The architecture incorporates a branching structure where data flows into distinct pathways for feature computation. By combining advanced components such as feature extraction blocks, residual layers, dense connections, and multi-filter utilization, our proposed model captures features at multiple scales, enabling it to recognize both fine-grained details and broader contextual information. Strategically placed residual layers further enhance the network's ability to extract fine-grained features from ulcer images, improving model transparency and interpretability.

This innovative design embodies a strategic fusion of image-based features, demonstrating its potential to revolutionize the accuracy and comprehensibility of DFU classification. By efficiently capturing complex features and ensuring computational efficiency, it has the potential to significantly contribute to improved patient care and clinical decision-making in diabetic foot ulcer management.

H. Proposed Dense Deep Neural Architecture powered by Generative Adversarial Networks

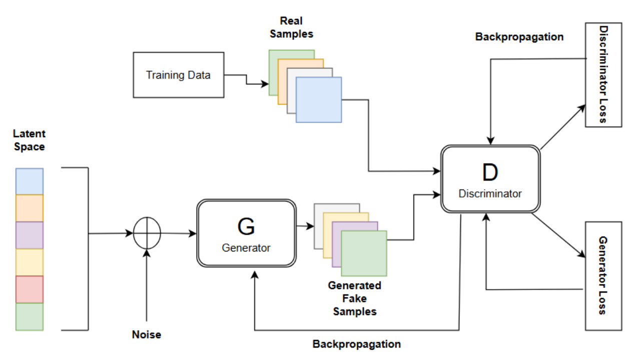

Fig. 2. GAN Architecture

A Generative Adversarial Network (GAN) is a deep learning framework consisting of two competing neural networks: the Generator and the Discriminator, trained in a game-theoretic manner. One uses the noise as input and generates samples. The second model which is known as discriminator receives samples from the generator and the training data. According to game theory, the generator is trained to produce an image that looks like a real image, whereas the discriminator is learning to discriminate perfectly from generated data to actual data. The proposed GAN architecture is designed for the purpose of data augmentation which overcomes the problem of class imbalance in our dataset. In DFU, Generative Adversarial networks are used for increasing the medical samples for improving the data quality and quantity both to get better classification performance. Generator: Creates realistic DFU images for data augmentation, covering underrepresented classes (e.g., infections or ischemia). Discriminator: Ensures the generated images resemble the real DFU dataset, helping the classification model generalize better. On the other hand, Convolutional Neural network is added with GAN to efficiently process DFU images for classification tasks. Its emphasis on deep convolutional layers, skip connections, and bottleneck layers contributes to its ability to capture complex features while maintaining computational efficiency.

1) Latent Space : Latent space refers to the input space from which the generator creates synthetic data. It is usually a multidimensional vector of random noise (e.g., Gaussian or uniform distribution). Encodes high-level, abstract representations of features. By sampling different vectors from this space, the generator can produce diverse outputs. It helps to map noise to realistic data.

2) Generator Module: The generator aims to create synthetic data that is indistinguishable from real data. It takes random noise (z) as input from the latent space and passes it through layers (e.g., fully connected, convolutional, or transposed convolutional layers). It then outputs synthetic data (e.g., an image, audio, or text). The objective of generator is fool the discriminator into classifying generated data as real. The generator is then optimized using the generator loss, which is based on the discriminator's feedback.

3) Discriminator Module: The discriminator acts as a classifier that distinguishes between real and generated data. It takes input from both real data (e.g., images from the dataset) and synthetic data (from the generator) and passes the data through layers (e.g., convolutional layers) and outputs a single value (real or fake). Its objective is to accurately classify real data as "real" and synthetic data as "fake." The discriminator is then optimized using the discriminator loss, which aims to maximize the classification accuracy during backpropagation.
Objective of GAN: The GAN framework can be thought of as a min-max game:Discriminator's Goal: Maximize classification accuracy. Generator's Goal: Minimize the discriminator's ability to distinguish real from fake data.

$$\min_G \max_D V(G, D) = \mathbb{E}_x[\log(D(x))] + \mathbb{E}_z[\log(1 - D(G(z)))] \tag{1}$$

4) Discriminator Loss : Measures the ability of the discriminator to correctly classify real and fake data. D(x) : Probability of real data being classified as real. 1-D(G(z): Probability of fake data being classified as fake. Maximize D(x) for real data and 1-D(G(z)) for fake data.

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^{m} \left[\log D\left(x^{(i)}\right) + \log\left(1 - D\left(G\left(z^{(i)}\right)\right)\right)\right] \tag{2}$$

5) Generator Loss : Measures how well the generator is fooling the discriminator. D(G(z)) : Probability of the discriminator classifying the generated data as real. Minimize the probability of the discriminator identifying generated data as fake.

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^{m} \log\left(1 - D\left(G\left(z^{(i)}\right)\right)\right) \tag{3}$$

6) Backpropagation: Backpropagation is used to minimize the loss functions for both the Generator and the Discriminator. It propagates the error signals (gradients) backward through the networks, allowing the weights to be updated using gradient descent. The process involves two main steps: Update the Discriminator: Improve its ability to distinguish between real and fake data. Update the Generator: Improve its ability to fool the Discriminator. The generator and discriminator are updated alternately, not simultaneously. The discriminator's gradients depend on real and fake data, while the generator's gradients depend on the discriminator's feedback. When updating the generator, backpropagation flows through both the discriminator and the generator because the generator's output is fed into the discriminator. By iteratively applying backpropagation, GANs learn to generate data that closely resembles the real data distribution.

7) Convolutional Layers: The architecture starts with a series of convolutional layers that are responsible for extracting features from the input image.The convolutional layers use small filters to scan the image, capturing patterns and local features.These layers are designed to detect simple edges and textures in the initial layers and progressively more complex features like shapes and object parts as the depth increases.

8) Residual Blocks: The proposed model heavily utilizes residual blocks, which are introduced to address the vanishing gradient problem in very deep networks. A residual block consists of multiple convolutional layers followed by a shortcut connection that skips one or more layers. This connection allows gradients to flow directly through the network, which facilitates the training of very deep architectures.

9) Skip Connections: This is a key feature of the proposed design. Skip connections allow the network to access features from earlier layers directly, helping retain fine-grained information from the early layers and combining it with high-level features from deeper layers. This strategy enhances the network's ability to

detect complex patterns at varying levels of abstraction, improving overall feature extraction and enabling more accurate classification.

10) Downsampling: Pooling layers and stridden convolutions are used for downsampling feature maps, reducing their dimensions. This downsampling process increases the receptive field of the later layers, allowing the network to capture larger context and higher-level features. Additionally, downsampling speeds up processing by reducing the computational load in subsequent layers, making the network more efficient.

11) Fusion: Fusion of features is employed at multiple stages, where high-level features are combined with lower-level features. This approach helps the network capture both local and global contextual information, enhancing its ability to understand complex patterns and improving classification performance.

12) Global Average Pooling: At the end of the architecture, global average pooling is applied to compute the average value of each feature map, resulting in a fixed-size vector. This vector contains aggregated information from the entire image, enabling the network to grasp the overall context and making it easier to perform classification.

13) Fully Connected Layers: The global average pooled vector is connected to a few fully connected layers, transforming it into a format suitable for the final classification task. These layers help in making final predictions by mapping the aggregated features to the output classes.A dense layer with sigmoid activation for binary classification is used in the architecture.

14) Activation Functions and Non-Linearity: Throughout the architecture, activation functions like Leaky ReLU (Rectified Linear Unit) are applied after each convolutional layer. These activation functions introduce non-linearity into the network, allowing it to learn complex relationships in the data and enabling the network to model intricate patterns more effectively.

15) Softmax Layer: The softmax layer is a crucial component in deep neural network models designed for classification tasks. It converts the raw scores or logits produced by the previous layer into normalized probabilities, indicating the likelihood of each class being the correct classification for the input data. This probabilistic output helps in making the final decision regarding the DFU classification.

I. Hyperparameter Setting and Optimization

Proper tuning of hyperparameters is essential to achieve optimal performance for the proposed DFU_GANNet architecture. Key hyperparameters include learning rate, batch size, number of epochs, activation functions, regularization techniques (e.g., dropout and L2 regularization), optimizers (e.g., Adam, SGD), and learning rate schedulers.
Some of the common hyperparameters that are used in deep model design and their requirements are given below:

1) Learning Rate (LR): Learning Rate Scheduler reduces the learning rate when validation accuracy plateaus, with a factor of 0.1 and patience of 3 epochs. Controls the step size during gradient descent optimization. A scheduler adjusts the rate dynamically to ensure efficient convergence without divergence. A learning rate of 0.001 is used.
2) Batch Size: Specifies the number of data samples processed per iteration. Balances computational efficiency and generalization, typically ranging from 16 to 64 depending on available memory.
3) Number of Epochs: Defines the number of complete passes through the dataset during training. Early stopping criteria prevent overfitting by halting training when validation performance stops improving.
4) Activation Functions: Non-linear transformations (e.g., ReLU, Leaky ReLU) applied after each layer to model complex relationships in data effectively.
5) Regularization: Techniques like dropout (e.g., 0.3–0.5) and L2 regularization (e.g., 0.0001–0.01) mitigate overfitting by controlling model complexity.
6) Optimizer: Algorithms like Adam with momentum are used to minimize the loss function efficiently, with adjustments based on specific task requirements. We have used Adam optimizer with a learning rate of 0.001.
7) Learning Rate Scheduler: Dynamically reduces the learning rate during training (e.g., decay by a factor of 0.1 every 30 epochs) to improve convergence and avoid overshooting minima.
8) Loss Function: Objective function guiding the optimization process. Sparse Categorical Cross-Entropy for multi-class classification is used for tasks like DFU classification.
9) Drop out: Dropout rate of 0.4 was used for regularization to mitigate overfitting.
10) Early Stopping: Monitors validation performance to halt training automatically when no further improvement is observed, avoiding overfitting. Halts training if validation accuracy does not improve for 4 consecutive epochs to prevent overfitting.

IV. Experimental Results

A. Dataset Description

To evaluate the performance of the proposed deep neural architecture during experimentation, we used a licensed dataset that consists of images of patient's foot with DFU over the previous five years at the Lancashire Teaching Hospitals, obtaining ethical approval from all relevant bodies and patients written informed consent. We have received dataset to use diabetic foot ulcer (DFU) images for the purpose of this research. These images are photographs collected from the Lancashire Teaching Hospitals, where photographs were acquired from the patients during their clinical visits. Three cameras were used for capturing the foot images, Kodak DX4530, Nikon D3300 and Nikon COOLPIX P100. The images were acquired with close-ups of the full foot at a distance of around 30–40 cm with the parallel orientation to the plane of an ulcer. The total images for DFUC2021 is 15,683, with 5,955 images for training set and 5,734 images for testing set. Part A has 1038 Abnormal images and Part B has 641 Normal images whereas Part B has 5890 images of Infection class and 9870 images of Ischaemia class.

TABLE I
DFUC2021 Dataset Description

Data Type	Internal Sub Division	Sample Count	Dimension
DFUC2021_train	Images	5955	224x224
DFUC2021_test	Images	5734	224x224
PartA_DFU_Dataset	Abnormal Normal	1038 641	Variable Variable
PartB_DFU_Dataset	Infection Ischaemia	5890 9870	256x256 256x256

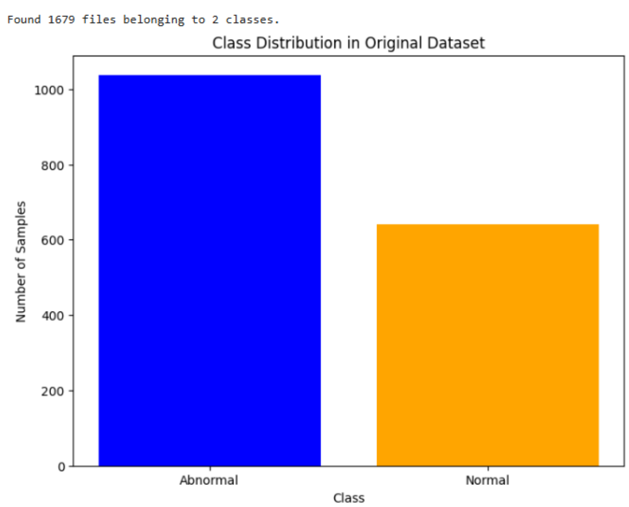

Fig. 3. Data Imbalance between Abnormal and Normal Class

B. Performance Visualization

Training and validation accuracy were plotted across epochs to assess model convergence. Predictions on both training and test datasets were visualized to qualitatively evaluate classification performance.

V. Conclusion and Future Work

TABLE II
DFUC2021_dataset: Impact of Data Augmentation on the Dataset

Data Type	Internal Sub Division	Samples Before	Samples After
PartA_DFU Dataset	Abnormal	1038	1038
	Normal	641	1038

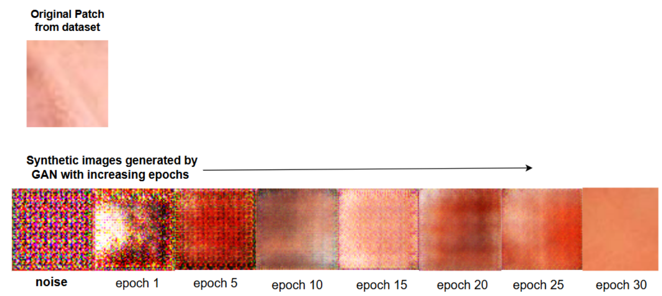

Fig. 4. GAN Generated Images

Diabetic foot ulcers (DFUs) are a critical complication of diabetes that pose significant challenges to patient health and healthcare systems. Addressing these challenges requires effective classification methodologies that can assist in timely diagnosis and treatment. Many studies explained that GAN is used to develop the synthesized images to diagnose diseases.In this study, we introduced a comprehensive approach to DFU classification, incorporating advanced deep learning techniques and novel architectural enhancements. By leveraging GAN-based data augmentation, we effectively mitigated class imbalance and improved data diversity, enabling robust training of our proposed deep neural network. Future targets include developing a mobile app for detecting and monitoring diabetic foot ulcers (DFUs) at home that can greatly help patients and their family members in detection and tracking of ulcers, reducing the need for frequent hospital visits. This would lower the risk of complications, especially for people with diabetes who are more vulnerable to serious illness. The app could also ease the pressure on healthcare systems by enabling timely care and reaching patients in remote areas, making it a valuable tool for better and safer DFU management. It may also include expanding the dataset, testing the model in real-world clinical settings, and integrating it into diagnostic tools for broader application in diabetic foot management, paving the way for improved patient care and aiding clinicians in making informed decisions.

References

[1] Nishu Bansal1, Ankit Vidyarthi1 "DFootNet: A Domain Adaptive Classification Framework for Diabetic Foot Ulcers Using Dense Neural Network Architecture", DOI: https://doi.org/10.1007/s12559-024-10282-4 Published on 29th April 2024. Springer Science+Business Media, LLC, part of Springer Nature 2024.

[2] Moi Hoon Yapa,*, Ryo Hachiumab, Azadeh Alavic, Raphael Brüngeld,g, Bill Cassidya, Manu Goyale, Hongtao Zhuf, Johannes Rückertd, Moshe Olshanskyc, Xiao Huangf, Hideo Saitob, Saeed Hassanpoure, Christoph M. Friedrichd,g, David B. Ascherc, Anping Songf, Hiroki Kajitah, David Gillespiea, Neil D. Reevesa, Joseph M. Pappachani, Claire O'Sheaj, Eibe Frank, "Deep learning in diabetic foot ulcers detection: A comprehensive evaluation" , DOI:

https://doi.org/10.1016/j.compbiomed.2021.104596 Received 17 March 2021; Received in revised form 17 June 2021; Accepted 17 June 2021

[3] LAITH ALZUBAIDI1,2, ALAA AHMED ABBOOD2, MOHAMMED A. FADHEL3, OMRAN AL-SHAMMA2, JINGLAN ZHANG1 1School of Computer Science, Queensland University of Technology, Brisbane, QLD 4000, Australia 2AlNidhal Campus, University of Information Technology and Communications, Baghdad 10001, Iraq 3 College of Computer Science and Information Technology, University of Sumer, Thi Qar 64005, Iraq *Corresponding Author: laith.alzubaidi@hdr.qut.edu.au, July 2021 , “Comparison of hybrid convolutional neural networks models for diabetic foot ulcer classification” ; Journal of Engineering Science and Technology.

[4] Laith Alzubaidi1, Mohammed A. Fadhel2, Sameer R. Oleiwi3, Omran Al-Shamma4, Jinglan Zhang5 , “DFU_QUTNet: Diabetic Foot Ulcer Classification Using Novel Deep Convolutional Neural Network”, DOI: https://doi.org/10.1007/s11042-019-07820-w, Multimedia Tools and Applications.

[5] Manu Goyala, Neil D. Reevesb, Satyan Rajbhandaric, Naseer Ahmadd, Chuan Wange, Moi Hoon Yapa, “Recognition of ischaemia and infection in diabetic foot ulcers: Dataset and techniques” DOI: https://doi.org/10.1016/j.compbiomed.2020.103616, Received 27 October 2019; Received in revised form 9 January 2020; Accepted 9 January 2020, Published by Elsevier Ltd.

[6] MOHAMMUD SHAAD ALLY TOOFANEE1,2, SABEENA DOWLUT2, MOHAMEDHAMROUN1,4, KARIM TAMINE1, VINCENT PETIT2, ANH KIET DUONG3 and DAMIEN SAUVERON1 1Department of Computer Science, XLIM, UMR CNRS 7252, University of Limoges, 123,Avenue Albert Thomas, 87060 Limoges, France 2Université des Mascareignes, Concorde Avenue Roches Brunes Rose Hill,Mauritius 3Faculty of Science and Technology, University of Limoges, 23,Avenue Albert Thomas, 87060 Limoges, France 43iL Ingénieurs, 43 rue de Sainte Anne, 87015 Limoges, France, “ DFU-SIAM a Novel Diabetic Foot Ulcer Classification with Deep Learning” , DOI 10.1109/ACCESS.2023.3312531, Author et al.: Preparation of Papers for IEEE TRANSACTIONS and JOURNALS.

[7] Abdul Qayyum1, Abdesslam BENZINOU1, Moona Mazher2, and Fabrice Meriau deau3 1ENIB, UMR CNRS 6285 LabSTICC, Brest, 29238, France 2 Department of Computer Engineering and Mathematics, University Rovira i Virgili, Spain 3ImViA Laboratory, University of Bourgogne Franche-Comté, Dijon, France qayyum@enib.fr , “Efficient Multi-Model Vision Transformer based on Feature Fusion for Classification of DFUC2021 Challenge”, January 2022 DOI: 10.1007/978-3-030-94907-5_5.

[8] SeungHwan Kim School of Artificial Intelligence Sungkyunkwan University Korea, Suwon shkim218@skku.edu , Sukhan Lee* School of Artificial Intelligence Sungkyunkwan University Korea,Suwon lsh1@skku.edu , ” Self-Supervised Augmentation of Quality Data Based on Classification-Reinforced GAN”, 2023 17th International Conference on Ubiquitous Information Management and Communication (IMCOM).

[9] Bill Cassidy1, Neil D. Reeves2, Joseph M. Pappachan3, David Gillespie1, Claire OShea4, Satyan Rajbhandari3, Arun G. Maiya5, Eibe Frank6, Andrew J.M. Boulton7, David G. Armstrong8, Bijan Naja 9, Justina Wu4, and Moi Hoon Yap1 1 Department of Computing and Mathematics, Faculty of Science and Engineering, Manchester Metropolitan University, Manchester, M1 5GD, UK 2 Research Centre for Musculoskeletal Science & Sports Medicine, Faculty of Science and Engineering, Manchester Metropolitan University, Manchester, M1 5GD, UK 3 Lancashire Teaching Hospital, Preston PR2 9HT, UK 4 Waikato District Health Board, New Zealand 5 Manipal College of Health Professions, India 6 Department of Computer Science, University of Waikato, New Zealand 7 School of Medical Sciences, University of Manchester, UK 8 Keck School of Medicine, University of Southern California, USA 9 Baylor College of Medicine, Texas, USA, “DFUC 2020: Analysis Towards Diabetic Foot Ulcer Detection”, arXiv:2004.11853v3 [cs.CV] 24 May 2021

[10] Moi Hoon Yap , Bill Cassidy , Joseph M. Pappachan , Claire O’Shea , David Gillespie and Neil D. Reeves Faculty of Science and Engineering, Manchester Metropolitan University, Manchester, UK Lanchashire Teaching Hospitals, UK Waikato District Health Board, Hamilton, New Zealand, “Analysis Towards Classification of Infection and Ischaemia of Diabetic Foot Ulcers”, arXiv:2104.03068v2 [cs.CV], Published: 21 June 2021

[11] Manu Goyal, Student Member, IEEE, Neil D. Reeves, Adrian K. Davison, Member, IEEE, Satyan Rajbhandari, Jennifer Spragg, Moi Hoon Yap, Member, IEEE, ”DFUNet: Convolutional Neural Networks for Diabetic Foot Ulcer Classification”, arXiv:1711.10448v2 [cs.CV], 10 Dec 2017