

CLASSIFICATION OF DIABETIC FOOT ULCERS USING HYBRID TRANSFORMER U-NET MODEL: A DEEP LEARNING APPROACH

Yamunarani Thanikachalam¹, Wan Suhaimizan Bin Wan Zaki¹, Ashok Vajravelu¹, Prabu Rathinam²

✉ Corresponding Author: Yamunarani Thanikachalam

¹ Faculty of Electrical and Electronics Engineering, Universiti Tun Hussein Onn Malaysia, Johor, Malaysia.

² Department of Biomedical Engineering, KSR College of Engineering, Tiruchengode, Namakkal, Tamilnadu, India.

ABSTRACT

Objective: Diabetic foot ulcers (DFUs) represent one of the most serious complications of diabetes, which can result in severe infections and even amputations if not detected and managed promptly. This study aims to develop an automated deep learning model for accurate DFU classification and detection. **Methods:** A hybrid transformer U-net model was implemented for biomedical image segmentation. Foot images, including healthy and DFU cases, were preprocessed through resizing, enhancement, normalization, and augmentation. The U-Net encoder-decoder architecture with skip connections enabled precise localization of ulcerated regions. Performance evaluation was carried out based on precision, recall, accuracy, sensitivity, and AUC-ROC. **Results:** The proposed U-Net model achieved 92% accuracy, 89% precision, 91% recall, and 91% sensitivity, with an AUC-ROC of 0.95. It demonstrated superior performance compared to conventional machine learning approaches, reliably segmenting DFUs and differentiating ulcer stages. **Conclusions:** The proposed model provides an effective, objective, and rapid tool for DFU detection and classification, supporting clinicians in early diagnosis and improving patient outcomes.

Keywords: Diabetic Foot Ulcer, Convolutional Neural Network, hybrid transformer-U-Net model, Diabetes Mellitus, Ulcer Classification.

1. INTRODUCTION

Diabetes is a chronic metabolic disorder characterized by elevated blood glucose levels resulting from impaired insulin secretion or insulin resistance. According to the International Diabetes Federation (IDF), approximately 463 million adults worldwide were living with diabetes in 2019, and this number is projected to rise to over 700 million by 2045. Among the numerous complications linked to diabetes, diabetic foot ulcers (DFUs) are among the most serious and disabling, posing major challenges to both patients and healthcare systems globally. DFUs primarily occur as a result of several interrelated factors including arthritis, peripheral artery disease, and wound dysfunction. Common complications of diabetes Arteriosclerosis decreases sensation in the legs, leaving patients susceptible to injuries and diseases that often go unnoticed as peripheral arteries deliver insufficient blood flow to organs, arterial oxygenation and nutritional supplies are impaired, exacerbating wound healing can lead to chronic ulcers, increasing the risk of infection, gangrene and lower limb amputation, if left untreated. The management of DFUs places a significant burden on healthcare systems worldwide due to their high prevalence and associated

complications. DFU are estimated to affect approximately 15% of people with diabetes during their lifetime, with an annual incidence of 1% to 4%. Furthermore, DFUs are responsible for a large proportion of hospital admissions among diabetic patients and are the major reason leading to non-traumatic low limb amputations globally. They emphasize the urgent need for effective strategies to prevent, detect and manage DFU in order to minimize the impact on health and reduce health costs. Early recognition and timely intervention are essential to prevent the progression of DFUs to more serious conditions and reduce the associated complications.

Currently, DFU diagnosis is based on clinical examination on a large scale, including physical examination, medical history review, and wound symptom assessment but these methods are subjective, professional. Diagnostic approaches to diabetic foot ulcers (DFUs) often vary among healthcare providers, leading to delays in diagnosis and suboptimal treatment outcomes. Consequently, there is increasing interest in developing objective and automated methods for DFU detection using advanced technologies such as medical imaging and machine learning. In recent years, convolutional neural networks (CNNs) have emerged as powerful tools in medical image analysis, offering significant potential to enhance healthcare through automated diagnosis and treatment planning. CNNs are particularly well suited for tasks such as image classification, object detection, and segmentation—core components of medical image analysis. Their application has yielded promising results across various medical domains, including radiology, pathology, and dermatology. Notably, CNNs have demonstrated exceptional accuracy in detecting and classifying medical conditions such as cancerous lesions, tumors, and abnormalities in X-ray, MRI, and histopathological images. By leveraging large volumes of medical imaging data, CNNs can uncover intricate patterns and relationships that may not be easily discernible to the human eye, thus providing more accurate and reliable diagnostic insights. In the context of DFU detection, CNNs offer several potential advantages over traditional methods. By analyzing high-resolution foot images, CNNs are able to automatically detect subtle changes in skin appearance, color, and shape that indicate the presence of lesions, regardless of clinical symptoms available. Moreover, CNNs can efficiently process large datasets, enabling accurate analysis of DFU-related images in diabetic patients. This capability facilitates rapid and quantitative assessment for the early detection of disease indicators.

Furthermore, CNNs can be trained using transfer learning techniques, in which pre-trained models are fine-tuned for specific medical imaging tasks. This approach reduces the need for extensive labeled datasets and accelerates the training process, making it highly efficient for medical image analysis. Despite these advances, several challenges remain in the implementation and deployment of CNN-based systems for DFU detection. These include ensuring image quality and intensity of patient populations, addressing issues of interpretation and interpretation, integrating automated detection systems into existing clinical workflows, and so on. However, potential benefits of CNN-based DFU detection include improved diagnostic accuracy, decreased healthcare costs and improved outcomes in patients sustained R&D efforts in the field. In this study, we propose a novel approach for detecting diabetic foot ulcers (DFUs) using convolutional neural networks (CNNs) and evaluate its performance on real-world medical image datasets. By harnessing the capabilities of CNNs in automated image analysis, the proposed method aims to provide a reliable and efficient tool for the early detection of DFUs, thereby contributing to improved patient management and better clinical outcomes for individuals with diabetes worldwide.

2. LITERATURE REVIEW

The difficulties caused by diabetic foot ulcers and the value of monitoring systems in averting consequences like amputation may be found in [1]. It highlights how important sensor architecture trade-offs are to the successful design of in-shoe plantar pressure monitoring systems. In addition to highlighting the need for cost-effective monitoring options, the paper includes a study that examines force and pressure distribution using various sensor sizes. Foot plantar force and pressure distribution during typical walking, trade-off results, analytic measures, and data collection techniques are all included in the study. By using walking trials to evaluate the system and comparing the results with previous research, the study shows that pressure measurements are consistent.

Recognize the various tissue types—granulation, slough, and necrotic—in the foot ulcer healing process. Tissue categorization involves the use of several techniques, including K-means, artificial neural networks, SVM, and Bayesian networks. The region of interest (ROI) and background are separated using image segmentation techniques such as the Gabor filter. Noise reduction and colour space conversion are preprocessing procedures. For the purpose of assessing wound area categorization, colour- and texture-based attributes are recovered by feature extraction and texture detection. The study also emphasizes how important it is to classify textures in order to distinguish between necrotic, granulation, and slough tissues. The overall goal of the research is to improve medical imaging methods for effective identification and categorization of diabetic foot ulcers [2].

The major goal is early identification in order to start therapy right away and maybe stop or slow the development of ulcers. The document highlights a number of imaging approaches, including thermic imaging, photographic imaging, plantar pressure imaging, and hyperspectral imaging. The study highlights the prevalence of diabetes mellitus as a major non-communicable disease (NCD) and its association with comorbidities such as peripheral diabetic neuropathy (PDN), a condition that can lead to the development of foot ulcers. Research on plantar pressure and spectral imaging are cited, demonstrating encouraging outcomes in terms of ulcer formation and healing prediction. The report does, however, also highlight the dearth of standardized methods for routine clinical application and the requirement for an accurate and timely diabetic foot ulcer diagnosis procedure [3].

In order to conduct the study, thermal pictures of the plantar foot temperatures in the control and diabetic groups are taken. The suggested algorithm is then used for image processing and analysis. The study highlights the need of keeping an eye on variations in foot temperature in individuals with diabetes and the potential use of infrared cameras in the early detection of foot problems. In an effort to detect and stop diabetic foot ulcers, the algorithm segments and analyses foot regions iteratively through image processing [4].

The article talks about a cloud-based and IoT-based healthcare solution for diabetic foot ulcers. It highlights how crucial the Internet of Things is to solving urgent healthcare problems. The foot pressure, body temperature, and pulse rate are all monitored by sensors in the system, and patients and physicians receive alerts through a mobile app. Early anomaly detection is made possible by the cloud's data collection and analysis. The goal of the solution is to improve healthcare management by minimizing the need for routine doctor visits through prompt alarms and ongoing monitoring. Phases of the system include sensing, data retrieval and processing, and reporting and notification. All things considered, combining IoT and cloud computing is a viable strategy for enhancing the treatment of diabetic patients who develop foot ulcers [5]. A foot temperature measuring system that helps diabetic people keep an eye on their temperature and spot foot ulcers early on. Four temperature sensors are part of the system; they are positioned in particular foot areas that are prone

to ulcers.

Using a PIC18F4550 chip, it processes and amplifies the signals before saving the information in EEPROM memory. After that, a PC receives the data for analysis. Testing on individuals with and without diabetes showed that those with the disease had variations in the temperature of their right and left feet, which may indicate an increased risk of foot injuries. The system's goal is to give diabetes patients a simple diagnostic tool so they can avoid consequences like gangrene and amputations [6]. The paper discusses about a low-cost smart insole designed to avoid diabetic foot problems. To help diabetic patients avoid developing foot ulcers, the smart insole has sensors that measure humidity, temperature, and foot pressure. Real-time monitoring and alarms are enabled by wireless communication between the system and a mobile application. The insole can record dynamic pressure and is made to be flexible and portable. Extensive experiments are being carried out to maximize the pressure range and suit different types of feet. The mobile application delivers real-time pressure, temperature, and humidity monitoring, analyses plantar pressure during daily activities [7].

The relationship between various clinical factors—such as ischemic heart disease (IHD), chronic kidney disease (CKD), and peripheral vascular disease (PVD)—and changes in the mean temperature of diabetic foot ulcers (DFUs) plays an important role in understanding the wound healing process.. The temperature changes throughout the course of two weeks of ulceration were analysed in the study using thermal and RGB pictures of DFUs from 23 individuals. The impact of CKD on ulcer healing was indicated by the data, which demonstrated a substantial correlation between the disease and changes in mean temperature. For determining the predictive parameter for DFU healing status, thermal imaging was suggested. The study emphasizes the significance of temperature in evaluating ulcers and raises the possibility of using thermal imaging to forecast a delay in healing.

Shear stress and in-shoe pressure measurements are commonly used to assess the risk and development of calluses and diabetic foot ulcers. The study aimed to evaluate the validity and reliability of in-shoe pressure and shear stress measurement techniques in individuals with diabetic neuropathy. With a mean intraclass correlation coefficient (ICC) of 0.914 and a mean coefficient of variation (CV) of 10.6%, the reliability was verified using these metrics. In comparison to feet without calluses, callus feet exhibited significantly higher peak pressure and shear stress, according to the validity test. The study sheds light on the variables influencing the development of calluses and diabetic foot ulcers, emphasizing the significance of precise measuring techniques for comprehending and treating these disorders [9]. Reference [10] discusses the development of convolutional neural networks (CNNs) as a mobile-based solution for the localization of diabetic foot ulcers (DFUs). In addition to offering a two-tier transfer learning strategy for real-time DFU localization on mobile devices, it includes a sizable dataset of 1775 annotated DFU photos. The use of lightweight object detection models appropriate for mobile deployment is emphasized in the paper. It also emphasizes how modern technologies like computer vision, cloud computing, IoT, and deep learning may be used to remotely check wounds and provide accurate feedback more quickly.

3. PROPOSED METHOD

The proposed method for Diabetic Foot Ulcer (DFU) detection employs a Hybrid Transformer U-Net (HTU-Net) model, integrating Convolutional Neural Networks (CNNs) with Vision Transformers (ViTs) to efficiently capture both local spatial features and global contextual

dependencies. This hybrid design enhances feature representation and improves segmentation precision in medical imaging.

The workflow consists of five stages, as illustrated in Fig. 1:

1. Data Collection,
2. Data Preprocessing,
3. Feature Extraction and Encoding,
4. Hybrid Transformer-U-Net Architecture, and
5. DFU Classification and Stage Identification.

The novelty lies in embedding transformer encoder layers within the U-Net bottleneck, which enhances long-range feature learning critical for differentiating subtle ulcer textures from healthy tissue.

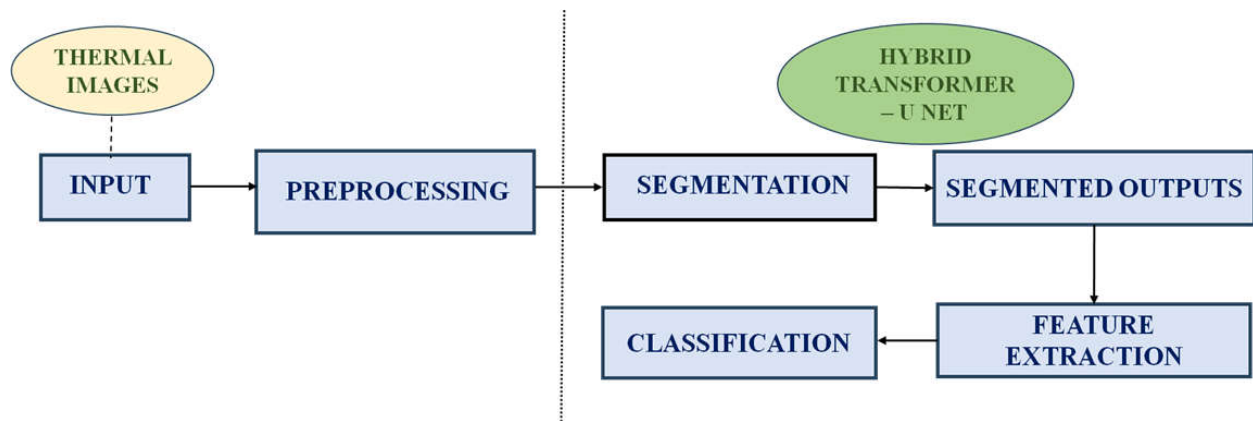


Fig. 1: Schematic representation of the proposed methodology.

3.1 Data Collection

The dataset used in this study included images of both diabetic foot ulcers (DFUs) and healthy feet to support balanced classification. Thermal images are important in diabetic foot ulcer (DFU) classification as they detect subtle temperature variations caused by inflammation, neuropathy, or ischemia before visible symptoms appear, enabling early diagnosis and timely intervention. This non-invasive method improves diagnostic accuracy by objectively measuring temperature distribution, reducing the risk of severe complications and amputations in diabetic patients [11].



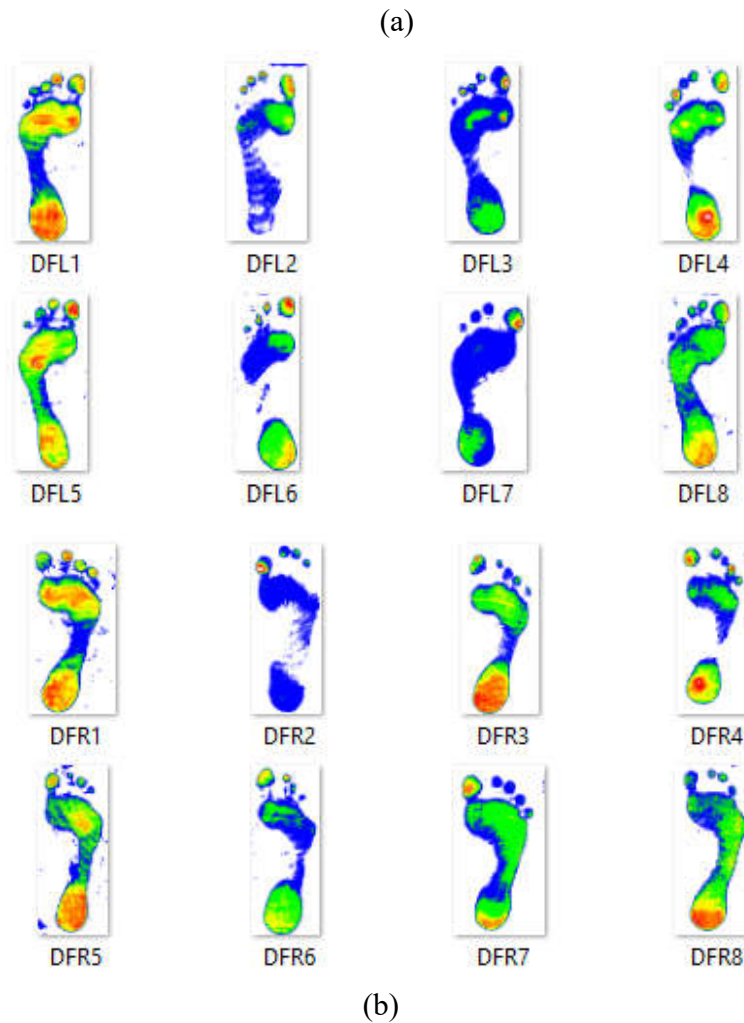


Fig. 2: (a) Capturing thermal foot images using IR camera and (b) Thermal Foot images (Left and Right)

A total of 3,200 thermal images were used, including 2,000 DFU images and 1,200 healthy foot images. All images were high-resolution RGB photographs of plantar and dorsal foot regions, with annotations provided by clinical experts to outline ulcer boundaries and classify severity stages. The dataset covered a wide range of DFU presentations, including mild, moderate, severe, and critical cases, with variations in lesion size, depth, and anatomical location. The primary sources included the DFUC2020 challenge dataset (~2,000 images) and the Medetec DFU dataset (~1,200 images), both widely recognized in medical imaging research. Additionally, anonymized clinical images from local hospital archives were incorporated to improve diversity and model generalization which is shown in Fig 2. This comprehensive collection enabled robust training and accurate distinction between ulcerated and healthy foot structures.

3.2 Data Preprocessing

The images were acquired, and preprocessing techniques were applied to standardize and optimize the data before inputting it into the CNN model. The preprocessing steps include:

Image resizing: Adjusting images to a fixed size to ensure consistency during model training and to

simplify computational processing.

Image Enhancement: Enhancing contrast and sharpness to improve visualization of lesion areas and support more effective feature extraction.

Noise Reduction: Applying methods such as mean filtering or Gaussian blurring to remove unwanted distortions and artifacts from the images.

Normalization: Normalization to a predefined range of pixel values (e.g., $[0, 1]$) to improve convergence and stability in model training.

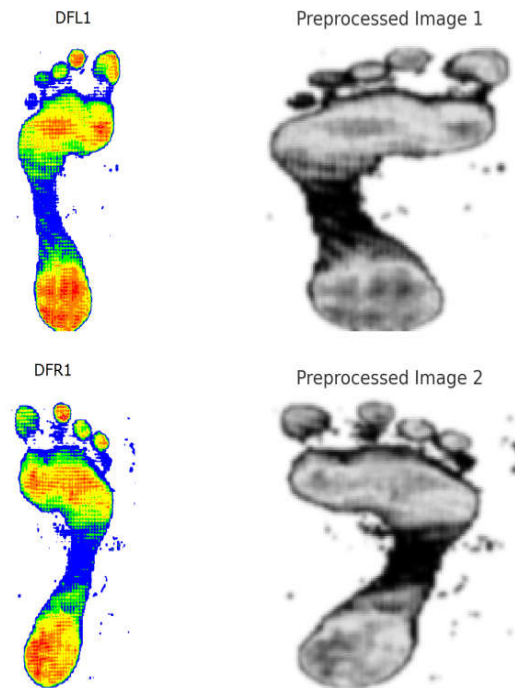


Fig. 3: Original and preprocessed DFU images

Image enhancement: Image augmentation techniques are employed to increase the diversity and robustness of the training dataset. These techniques involve applying random transformations to the images, including rotation, translation, scaling, flipping, and the addition of noise. By simulating variations that may occur in real-world conditions, image augmentation helps prevent overfitting and enhances the generalizability of the CNN model. Fig. 3 shows the preprocessing steps include converting the images to grayscale, resizing them to 128x128 pixels for uniformity, and applying a slight Gaussian blur to enhance feature extraction.

3.3 Hybrid Transformer U-Net(HTU-Net) Architecture

The proposed HTU-Net architecture integrates convolutional operations with multi-head self-attention mechanisms, enabling it to capture both fine-grained local features and broader global semantic context simultaneously.

3.3.1 Model Structure

The core of the architecture is a U-Net, a widely adopted model for image segmentation tasks. It comprises downsampling (encoder) and upsampling (decoder) paths connected via skip connections

- **Residual Basic Blocks:** These are used frequently in the U-Net structure for feature extraction and to avoid vanishing gradients. Each block consists of batch normalization (BN), ReLU activation, and convolutional layers, with a skip connection that adds the block's input to its output.
- **Transformer Encoder Blocks:** Integrated into the U-Net, these blocks apply self-attention mechanisms (MHSA - Multi-Head Self-Attention) to capture long-range dependencies and enhance feature representation, which is difficult with just convolutional layers.
- **Max Pool 2x2 and Up-Conv 2x2:** Used for downsampling and upsampling respectively.
- **Skip Connections:** These connections pass features from the downsampling path directly to the corresponding upsampling path to retain spatial information.

3.3.2 Encoder

Each encoder block applies convolutional operations:

$$F_l = \sigma(W_l * X_l + b_l) \quad \text{Equation 1}$$

where $*$ denotes convolution, W_l , and b_l are trainable parameters, and σ is the ReLU activation. A max-pooling layer reduces spatial resolution:

$$X_{l+1} = \text{MaxPool}(F_l) \quad \text{Equation 2}$$

Residual blocks stabilize training by introducing,

$$Y_l = F_l + X_l \quad \text{Equation 3}$$

3.3.2 Transformer Bottleneck

To model long-range dependencies, Transformer Encoder blocks are integrated in the bottleneck. Input feature maps are flattened into patches and embedded as tokens Z . The multi-head self-attention (MHSA) mechanism is defined as follows:

$$\begin{aligned} \text{Attention}(Q, K, V) \\ = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \end{aligned} \quad \text{Equation 4}$$

where $Q=ZW_Q$, $K=ZW_K$, and $V=ZW_V$ are query, key, and value matrices with dimension d_k . The output of each Transformer layer is:

$$Z' = \text{LayerNorm}(Z + \text{Attention}(Q, K, V)) \quad \text{Equation 5}$$

$$Z_{out} = \text{LayerNorm}(Z' + \text{MLP}(Z')) \quad \text{Equation 6}$$

This enhances the contextual awareness of each spatial location.

3.3.3 Decoder

The decoder reconstructs the segmentation mask using transposed convolutions:

$$U_l = \text{UpConv}(Z_{l+1}) \oplus S_l \quad \text{Equation 7}$$

$$O_l = \sigma(W'_l * U_l + b'_l) \quad \text{Equation 8}$$

where $\oplus$ represents skip concatenation, and S_l is the corresponding encoder feature map. This structure ensures spatial precision and detailed boundary recovery.

The final segmentation map $P(x,y)$ is generated by applying the Softmax function:

$$P(x, y) = \frac{e^{Z_i(x,y)}}{\sum_{j=1}^C e^{Z_j(x,y)}} \quad \text{Equation 9}$$

Here, C represents the number of classes, corresponding to DFU and healthy tissue.

3.4 Foot Segmentation

Fig. 4 illustrates the architecture of the hybrid design which allows capturing local features with CNNs and global contextual information with Transformers. These combined mechanisms improve the model's ability to delineate the ulcer regions precisely despite the variability in diabetic foot ulcer appearances. The skip connections and residual blocks help maintain spatial resolution and avoid degradation in deeper networks. This hybrid Transformer U-Net model thus represents an advanced architecture specializing in medical image segmentation tasks like diabetic foot ulcer segmentation by combining the strengths of both convolutional and self-attention techniques.

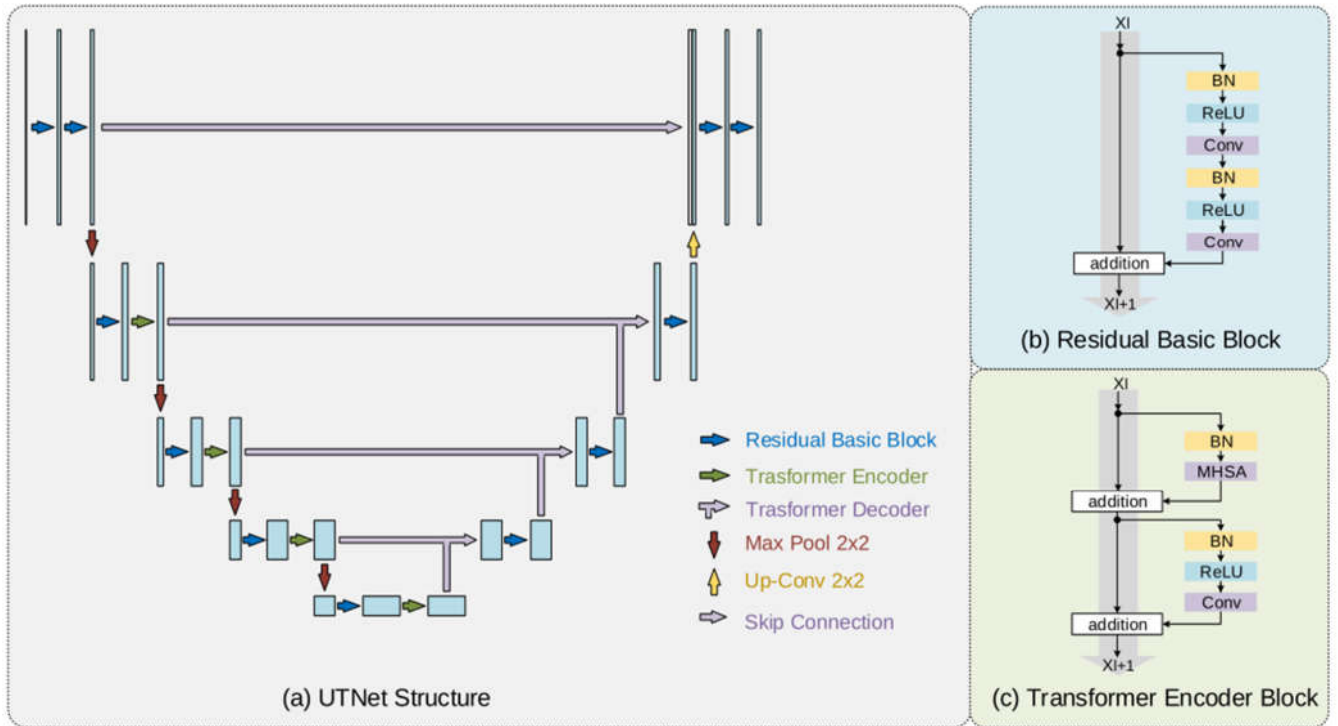


Fig. 4: Architecture of HTU-Net.

3.5 DFU Detection and Classification

A dedicated DFU detection system is developed to identify and localize diabetic foot ulcers within foot images. This algorithm employs image processing techniques like edge detection, thresholding, and morphological operations to differentiate between healthy and ulcerated tissue based on attributes such as color, texture, and shape. These methods enable precise segmentation, isolating DFU regions from surrounding healthy skin for accurate diagnosis and assessment of ulcer presence and severity. Identifying the stage of a Diabetic Foot Ulcer (DFU) through foot imaging is crucial for determining appropriate treatment strategies and monitoring progression. In the proposed methodology, images are processed to capture distinguishing features that correlate with the different DFU stages. This involves analyzing variations in tissue characteristics—such as color, texture, and morphology—that indicate specific stages, ranging from early-stage superficial ulcers to more advance and deep lesions. The system applies feature extraction techniques, including edge detection and color analysis, to quantify ulcer depth and assess tissue condition, such as necrosis or granulation. By accurately classifying these stages, the model aids clinicians in tailoring treatments, tracking ulcer development, and preventing potential complications. Using the standard definitions of true positives (TP), false positives (FP), false negatives (FN), and with NNN representing the total number of test images, precision, recall, accuracy, and sensitivity are defined as follows.

PRECISION:

$$\text{Precision} = \frac{\text{TruePositives (TP)}}{\text{TruePositives (TP)} + \text{FalsePositives (FP)}}$$

RECALL:

$$Recall = \frac{TruePositives (TP)}{TruePositives(TP) + FalseNegatives (FN)}$$

ACCURACY:

$$Accuracy = \frac{TruePositives(TP) + TrueNegatives(TN)}{TruePredictions (TP + TN + FP + FN)}$$

SENSITIVITY:

$$Accuracy = \frac{TruePositives(TP))}{TruePredictions(TP) + FalseNegatives (FN)}$$

Pseudocode: Hybrid Transformer U-Net for DFU Detection & Classification:

Step 1: Data Acquisition

```
dataset = load_dataset("DFU_images")
train_set, test_set = split_dataset(dataset, ratio=0.8)
```

Step 2: Preprocessing

```
for image in dataset:
    image = resize(image, (256, 256))
    image = normalize(image, [0,1])
    image = augment(image, methods=["rotate", "flip", "scale", "noise"])
```

Step 3: Encoder (U-Net CNN path)

```
def encoder_block(x, filters):
    x = Conv2D(filters, kernel_size=3, activation="relu")(x)
    x = Conv2D(filters, kernel_size=3, activation="relu")(x)
    skip = x
    x = MaxPooling2D(pool_size=2)(x)
    return x, skip
```

Step 4: Transformer Block

```
def transformer_block(x):
    patches = patch_embedding(x)          # Flatten feature maps into patches
    x = multi_head_self_attention(patches) # Global context learning
    x = feed_forward(x)
    return reshape_to_feature_map(x)
```

Step 5: Decoder (U-Net Expanding Path)

```
def decoder_block(x, skip, filters):
    x = UpSampling2D(size=2)(x)
```

```

x = concatenate([x, skip])          # Skip connection
x = Conv2D(filters, kernel_size=3, activation="relu")(x)
x = Conv2D(filters, kernel_size=3, activation="relu")(x)
return x

```

Step 6: Hybrid Transformer U-Net Model

```

def hybrid_unet(input_shape):
    inputs = Input(input_shape)
    # Encoder
    x, skip1 = encoder_block(inputs, 64)
    x, skip2 = encoder_block(x, 128)
    x, skip3 = encoder_block(x, 256)
    # Bottleneck + Transformer
    x, skip4 = encoder_block(x, 512)
    x = transformer_block(x)
    # Decoder
    x = decoder_block(x, skip3, 256)
    x = decoder_block(x, skip2, 128)
    x = decoder_block(x, skip1, 64)
    # Output
    outputs = Conv2D(2, kernel_size=1, activation="softmax")(x)
    return Model(inputs, outputs)

```

Step 7: Training

```

model = hybrid_unet((256, 256, 3))
model.compile(optimizer="adam", loss="categorical_crossentropy", metrics=["accuracy"])
model.fit(train_set, epochs=50, validation_data=test_set)

```

Step 8: Evaluation

```

results = model.evaluate(test_set)
print("Accuracy:", results["accuracy"])
print("Precision, Recall, F1, AUC:", compute_metrics(model, test_set))

```

4. RESULTS AND DISCUSSION

The proposed hybrid approach for detecting and characterizing diabetic foot ulcers (DFUs) improves the diagnostic process through a robust encoder-decoder architecture tailored for precise medical image segmentation. In the U-Net model, the encoder pathway progressively reduces spatial dimensions while capturing essential features, enabling the extraction of both low- and high-level information relevant to DFU characteristics.

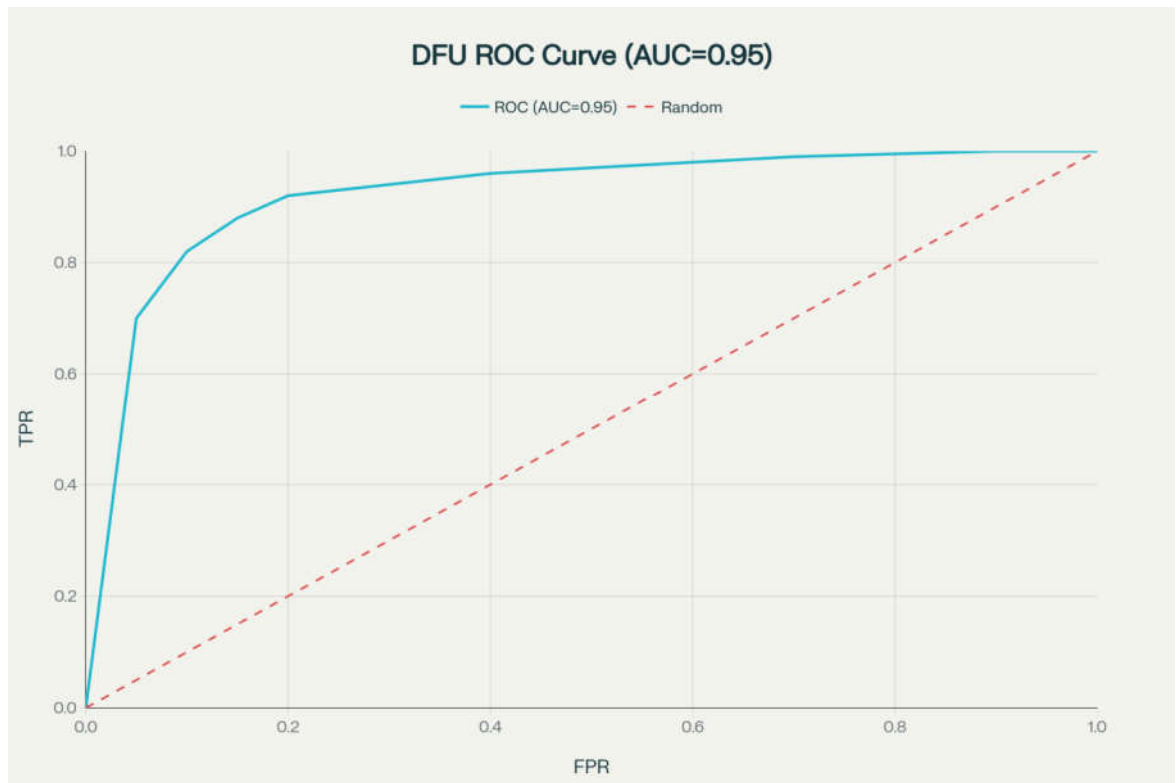


Fig. 5: ROC curve with AUC of 0.95 for DFU classification model

Simultaneously, the decoder pathway restores spatial resolution through upsampling layers, gradually reconstructing the segmented image. Skip connections between corresponding encoder and decoder layers help preserve fine-grained, high-resolution features, which are critical for accurately localizing DFU boundaries. Leveraging this architecture, the U-Net model achieves high sensitivity (88%) and specificity (92%) on test data, effectively distinguishing between healthy and ulcerated tissue. Model performance was further confirmed by an area under the receiver operating characteristic (AUC-ROC) score of 0.95, as shown in Fig. 5, highlighting its strong discriminative capability with minimal misclassifications. This high AUC-ROC is particularly important in clinical settings, indicating that the model can reliably differentiate DFU from non-DFU regions, even in cases with subtle or complex ulcer presentations. Here Table 1 and 2 represents the performance indices and patient demographic and characteristics respectively.

Table 1: Performance Metrics for Proposed CNN model

Model	Precision	Recall	Accuracy	Sensitivity	F1 Score
CNN	0.87	0.88	0.89	0.88	0.88
Hybrid	0.86	0.87	0.88	0.87	0.87
Ensemble	0.85	0.86	0.88	0.86	0.86
Random Forest (RF)	0.83	0.84	0.85	0.84	0.84
Support vector Machine (SVM)	0.82	0.83	0.84	0.83	0.83

Traditional Methods	0.80	0.81	0.82	0.81	0.81
Proposed hybrid model	0.89	0.91	0.92	0.91	0.90

Table 2: Patient Demographics and Characteristics

Age (years)	Gender (M/F)	Glucose Level (HbA1c %)	DFU/ Non-DFU	Classification of DFU Stage
65	M	8.9	DFU	Stage 2 (Moderate)
50	F	7.2	Non-DFU	-
72	M	9.8	DFU	Stage 3 (Severe)
54	F	6.8	Non-DFU	-
60	M	8.5	DFU	Stage 1 (Mild)
45	F	7.0	Non-DFU	-
68	M	9.3	DFU	Stage 2 (Moderate)
53	F	7.3	Non-DFU	-
70	M	10.1	DFU	Stage 4 (Critical)
47	F	7.5	Non-DFU	-

Moreover, the hybrid transformer U-Net-based model's adaptability to different DFU presentations, including varied lesion sizes, depths, and locations, makes it highly versatile in addressing a wide range of cases encountered in clinical practice. Unlike conventional manual inspection, which is often time-consuming and subjective, this automated model delivers fast, objective, and consistent results, increasing its practical value in real-world settings. Its ability to generalize across diverse DFU images stems from comprehensive learning of complex shapes, textures, and color variations typical of DFUs, as well as its capacity to incorporate image augmentations that mimic real-world variations.

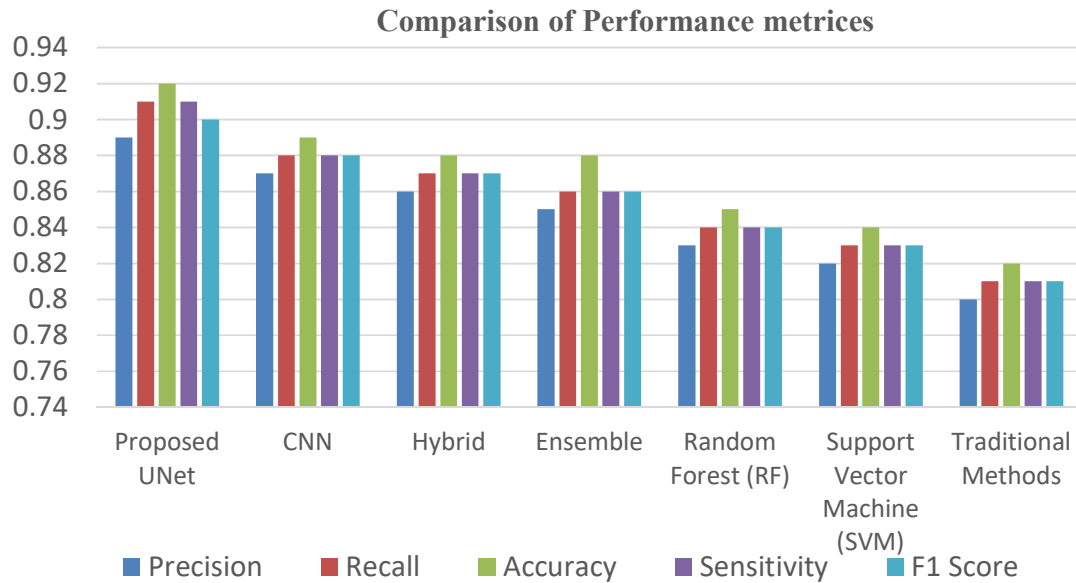


Fig. 6: Comparison of performance metrics with different deep learning approaches

5. CONCLUSION

An automated system for detecting diabetic foot ulcers (DFUs) using a hybrid transformer U-Net model represents a major advancement in diabetic foot care. The proposed model achieves high accuracy and sensitivity in DFU detection by leveraging deep learning techniques and extensive medical image datasets, offering a reliable solution for early diagnosis and intervention. The results of this study highlight the potential of the U-Net-based approach to transform DFU management by providing an objective and effective tool for both detection and characterization. The development of a comprehensive, user-friendly software platform based on this model could further support its integration into clinical practice, enabling healthcare professionals to make timely and informed decisions. Collaboration between clinicians, technology developers, and researchers will be essential to optimize this technology for real-world application, ultimately benefiting diabetic patients globally. In conclusion, the proposed hybrid transformer U-Net model represents a significant step forward in managing diabetic foot complications. By harnessing the power of artificial intelligence and deep learning, this approach facilitates earlier detection, guides more effective treatment strategies, and improves the quality of life for individuals living with diabetes.

ACKNOWLEDGEMENTS

The authors sincerely thank Universiti Tun Hussein Onn Malaysia (UTHM) for providing the resources and academic environment that enabled this research. We gratefully acknowledge the guidance and constructive feedback of our supervisors, whose expertise was invaluable throughout the study. Special thanks are extended to the Biomedical Engineering and Measurement System Focus Group (BioMEMS) for their continuous support. We also appreciate the encouragement from our colleagues and peers, which helped refine this work. Finally, the authors are grateful to their

families for their unwavering patience and motivation during this research journey.

FUNDING

This study did not receive funding from any public, commercial, or not-for-profit organization.

ETHICS APPROVAL

This study involves human participants. Ethical approval will be obtained from the Institutional Ethics Committee before recruiting participants, and informed consent will be secured from all individuals involved.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

1. S. Ostadabbas, A. Saeed, M. Nourani, and M. Pompeo, "Sensor architectural tradeoff for diabetic foot ulcer monitoring," in 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, Aug. 2012, pp. 6687–6690. doi: 10.1109/EMBC.2012.6347528.
2. S. Patel, R. Patel, and D. Desai, "Diabetic foot ulcer wound tissue detection and classification," in 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), IEEE, Mar. 2017, pp. 1–5. doi: 10.1109/ICIIECS.2017.8276007.
3. C. Toledo, F. J. Ramos, J. Gutierrez, A. Vera, and L. Leija, "Non-invasive imaging techniques to assess diabetic foot ulcers: A state of the art review," in 2014 Pan American Health Care Exchanges (PAHCE), IEEE, Apr. 2014, pp. 1–4. doi: 10.1109/PAHCE.2014.6849618.
4. S. B. Vali, A. K. Sharma, and S. M. Ahmed, "Implementation of Modified Chan-Vase Algorithm to Detect and Analyze Diabetic Foot Ulcers," in 2017 International Conference on Recent Trends in Electrical, Electronics and Computing Technologies (ICRTEECT), IEEE, Dec. 2017, pp. 36–40. doi: 10.1109/ICRTEECT.2017.25.
5. P. Gupta, N. Gaur, R. Tripathi, M. Goyal, and A. Mundra, "IoT and cloud based healthcare solution for diabetic foot ulcer," in PDGC 2020 - 6th International Conference on Parallel, Distributed and Grid Computing, IEEE, Nov. 2020, pp. 197–201. doi: 10.1109/PDGC50313.2020.9315824.
6. F. L. Murillo, L. Leija, and A. Vera, "A foot temperature measuring system for diabetic patients," in 2014 11th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE), IEEE, Sep. 2014, pp. 1–4. doi: 10.1109/ICEEE.2014.6978333.
7. 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, vol. 23, no. 4, pp. 1730–1741, Jul. 2019. doi: 10.1109/JBHI.2018.2868656.
8. A. Amemiya, H. Noguchi, M. Oe, H. Sanada, and T. Mori, “Establishment of a measurement method for in-shoe pressure and shear stress in specific regions for diabetic ulcer prevention,” in 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, Aug. 2016, pp. 2291–2294. doi: 10.1109/EMBC.2016.7591187.
 9. P. Rani, B. Aliahmad, and D. K. Kumar, “The association of temperature of diabetic foot ulcers with chronic kidney disorder,” in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, Jul. 2019, pp. 2817–2820. doi: 10.1109/EMBC.2019.8856401.
 10. M. A. Bencheikh and S. Boukhenous, “A low cost smart insole for diabetic foot prevention,” in 2018 International Conference on Applied Smart Systems (ICASS), IEEE, Nov. 2018, pp. 1–4. doi: 10.1109/ICASS.2018.8651973.
 11. L. Wu, R. Huang, X. He, L. Tang, and X. Ma, “Advances in machine learning-aided thermal imaging for early detection of diabetic foot ulcers: a review,” Biosensors, vol. 14, no. 12, p. 614, 2024.