Journal of Advanced Health Informatics Research (JAHIR)
Vol. 2, No. 3, 2024, pp. 146-167
ISSN: 2985-6124, DOI: 10.59247/jahir.v2i3.299

Potential Use of U-Net and Fuzzy Logic in Diabetic Foot Ulcer Segmentation: A Comprehensive Review

Rachman Hidayat [1], Annastasya Nabila Elsa Wulandari [2], Purwono Purwono [3], Khoirun Nisa [4]
[1-4] *Department of Informatics, Universitas Harapan Bangsa, Indonesia*

ARTICLE INFO

Article history:

Received May 09, 2024
Revised December 30, 2024
Published January 04, 2025

Keywords:

Diabetic Foot Ulcer;
U-Net;
Medical;
Image;
Segmentation;

ABSTRACT

Diabetic foot ulcer (DFU) image segmentation is still an interesting concern of researchers. Various new deep learning-based methods have been proposed to handle this image segmentation problem. Some research problems that are still faced by many researchers are dataset problems that are considered limited and need further clinical trials. The challenges of data problems include heterogeneity and image quality variations in the shape of skin lesions and subjectivity when annotating. The evaluation results from previous studies also show a considerable difference where there are still low accuracy results, but also too high accuracy is still found so that it is considered to have the potential for overfitting. As a result of the review of various related studies, there is an interesting potential of applying fuzzy logic to the U-Net architecture. This architecture has become very popular because it is widely used in medical image segmentation. The application of fuzzy logic can be applied to the U-Net architecture such as encoder, decoder, skip connection to adjust various U-Net parameters.

Corresponding Author:

Rachman Hidayat, Department of Informatics, Universitas Harapan Bangsa, Indonesia
Email: rachmanhidayat2906@gmail.com

1. INTRODUCTION

Diabetic foot ulcer is one of the more serious vascular complications of diabetes, and has affected 6.3% of the world's human population health [1]. Diabetic foot ulcers are foot infections, ulcers, and deep tissue damage caused by nerve abnormalities and varying degrees of vasculopathy in the distal lower extremities of diabetic patients [2]. The most characteristic clinical symptom is chronic ulceration, which can lead to toe amputation/amputation or even death. It has been claimed that the lifetime risk of developing a foot ulcer reaches 25%, with an annual incidence of 2-7% [3].

Regular check-ups and thorough treatment are essential to prevent more severe wounds [4]. Wound management requires regular wound assessment, documentation and treatment by medical professionals [5]. Trained nurses annotate relevant information in assessment reports which are then used to track healing progress and plan the most appropriate treatment strategy. Assessment is usually done by visual inspection of wound features, such as measurement of area and depth, and annotation of growth tissue [6]. Despite the importance of objective and accurate wound documentation, assessment reports are often inconsistent and sparse.

Wound imaging via cameras and smart phones has been adopted to support wound documentation with more objective tools [7]. Images of the wound are taken at each examination and medical professionals gain insight into the progress of wound healing by visually comparing with images taken at previous appointments.

Image analysis techniques can potentially be used to obtain all the necessary information for more objective wound documentation. One of the challenging tasks is semantic segmentation of medical images. Each image pixel is classified into foreground or background classes. In the case of wound images, accurate distinction between wound bed and background information is important as the wound bed area is considered one of the most relevant biomarkers for predicting wound healing [8] and excluding uninformative pixels can improve further analysis.

Image segmentation is one of the main tasks for medical image analysis needs and as a diagnostic support tool for health workers [9]. Accurate segmentation is essential in the clinical application of disease diagnosis and treatment planning [10]. Medical image segmentation has been used to identify clinically valuable lesions in medical images. For example, accurate vessel boundaries, definite lesion areas, and complete anatomical segmentation are very important for the clinical assessment of various diseases, including glaucoma, diabetic retinopathy, prostate, polyps, brain tumours, and other diseases [11].

Medical image segmentation is the process of dividing each pixel in an image into appropriate categories [12]. It is one of the most important steps in medical image analysis. Medical image segmentation helps extract detailed information from various tissues, organs, pathologies and biological structures for medical diagnosis, surgical planning and treatment.

In recent years, segmentation methods based on deep learning have gradually attracted wide attention due to their excellent segmentation performance. Deep learning can generalise the sample distribution and automatically extract target features from a large amount of data, which reduces the difficulty of manually setting image features [13]. Recently, deep neural networks have been used to build various medical image segmentation models. New models such as U-Net, DeepLab V3+, SegNet, RefineNet, and PSPNet have achieved accurate segmentation of various medical images for clinical diagnosis [11].

One of the popular semantic segmentation methods, U-Net has been widely used to automatically extract organ or lesion contours to overcome the limitations of manual segmentation [14]. U-Net is one of the most successful network architectures in medical image segmentation [15]. U-Net consists of two paths: contraction path (encoder or analysis path) and expansion path (decoder or synthesis path). The contraction path is mainly used to capture the context of the image, while the expansion path, which is symmetrical to the contraction path, is used to expand to the size of the original image [16].

U-Net has also been widely studied in the segmentation of diabetic foot ulcer images. In practice, there are still various weaknesses faced by research using U-Net in this field such as limited training data used [17] [18] and not representing all wound patients [19]. Other problems faced are the evaluation of segmentation models that still need to be improved in terms of accuracy [19] [20] to problems with the scalability of different devices [17].

Fuzzy logic is present to be one of the solutions in handling data uncertainty in artificial intelligence in the medical field [21]. Fuzzy logic is widely used in fuzzy medical image edge detection [22] [23]. In addition, there are also studies that utilise fuzzy logic in the main layer of deep learning such as fuzzy max-pooling layer [24] and fuzzy activation function layer [24] and fuzzy activation function layer [24]. Some of these findings are in line with what was developed by [25] which is using fuzzy layers in optimising manual thresholding, adaptive thresholding, gaussian thresholding, and otsu thresholding in medical image segmentation.

Based on some of the problems faced by many researchers in using U-Net in the field of medical image segmentation, especially diabetic foot wound image segmentation, there is a new research opportunity in the form of a fuzzy logic approach to U-Net architecture. The application of fuzzy logic to the U-Net architecture in the context of medical image segmentation has the opportunity to make a significant contribution. Fuzzy logic is an effective method to handle uncertainty and complexity in medical image data.

2. WAGNER SCALE OF DIABETIC FOOT ULCER

The Wagner scale is used to evaluate various types of wounds, especially those caused by vascular disorders in the leg and knee regions. One type of wound of particular concern is diabetic ulcers. Although the scale is currently applied to a wide range of wound types, its general use is more specifically directed at Diabetic Foot Ulcers [26].

Wound severity classification is important in medical care decision making especially in diabetic foot ulcer patients [27]. The Wagner scale provides a structure for determining wound severity by examining depth and breadth [26]. Wagner's classification system was developed in 1970 and has six levels of lesions from 0 to 5 degrees which can be seen in Table 1 [28] [29] .

Based on Table 1, Grade 0 is a pre-ulcerative/high-risk foot with no skin damage. Grade 1 is a superficial ulcer involving the epidermis, dermis or subcutaneous tissue. Grade 2 is a deeper wound that extends to the tendon, bone or joint and penetrates the subcutaneous tissue. Grade 3 is a deep wound with abscess/osteomyelitis and becomes as deep as grade 2 but with infection. Grade 4 is forefoot gangrene, requiring at least partial amputation. Grade 5 is gangrene of the entire leg and requires at least below-knee amputation.

Table 1. Wagner Scale

Grade	Information
Grade 0	Intact skin
Grade 1	Superficial ulcer
Grade 2	Deep ulcer
Grade 3	Ulcer with bone involvement
Grade 4	Forefoot gangrene
Grade 5	Full-foot gangrene

3. DATA SOURCES

Datasets used to identify diabetic foot ulcers can come from primary or secondary datasets. Many studies mention that these *datasets* have limitations that can affect the performance of the study [17] [30] [31] [19] [32]. This is certainly a challenge for researchers who are interested in handling this data.

Based on the results of the *literature review* that has been carried out, we found several types of datasets that are public, the rest are datasets created independently by researchers. Table 2 is some types of datasets that we have managed to collect information about.

Table 2. Data Availability

Dataset	Amount of Data	Type
DFUC2020 [32] [33]	4000	Public
Medetec [33]	-	Public
Second Intention Healing [33]	-	Public
Foot Ulcer Segmentation Challenge (FUSC) [33]	1000	Public
ESCALE database [34]	92	Public
Diabetic Wound Dataset [35]	1639	Primary
Extensive Database [17]	1775	Public
Taichung Rongmin General Hospital [36]	727	Primary
Annotated Wound Image [37]	1109	Primary
MICCAI 2021 Dataset [38]	1000	Public
DFU Dataset [39]	705	Primary
DFU Dataset [40]	844	Primary

4. CHALLENGES OF DIABETIC FOOT ULCER SEGMENTATION

Image segmentation in the context of Diabetic Foot Ulcers (DFUs) can be a challenging task due to various factors [41]. Diabetic foot ulcers are a common complication of diabetes, and accurately categorising ulcers from medical images is essential for diagnosis and treatment [18]. The following are some of the challenges associated with DFU image segmentation.

4.1 Varied Image Quality

Medical image quality may vary due to factors such as lighting conditions, differences in imaging equipment, and patient movement during image acquisition [42] [43]. Inconsistent image quality can make it difficult to apply standard segmentation techniques.

4.2 Diverse Ulcer Characteristics

DFUs can vary in size, shape, colour, and texture. Some ulcers may be superficial, while others may be deep [42], making it difficult to develop a one-size-fits-all segmentation algorithm.

4.3 Presence of Noise

Medical images often contain noise, artefacts or other distortions that can interfere with accurate segmentation [44]. Noise reduction techniques may be required to improve the reliability of the segmentation results [45]. Image noise reduction is an important task in the field of computer vision and image processing [46].

4.4 Complex Background

The foot region in medical images may have a complex background, including irregular shapes, uneven lighting, and skin colour variations [47]. For machine vision detection and identification under random texture complex background, the random texture complex background needs to be removed to extract the object without affecting the original features of the object [48].

4.5 Limited Annotated Data

One of the main problems inherent in medical applications for deep learning is that usually the amount of annotated data is limited [49]. Training an accurate segmentation model requires a large amount of annotated data [50]. However, obtaining annotated medical images, especially for rare conditions such as certain types of DFUs, can be challenging [19]. Limited data may lead to over or under generalisation [17].

4.6 Class Imbalance

The number of positive cases of ulcers in medical images is often much smaller than the negative cases (imbalance) [51]. This class imbalance can impact the performance of the segmentation model and may require special attention during training [52]. Failure to address this imbalance may lead to biased predictions, where the model favors the majority class while failing to accurately segment ulcer regions.

4.7 Computational Resource Requirements

Deep learning models, which are often used for image segmentation, are computationally intensive. Implementing and training these models may require large computational resources. Deep learning models, which are often used for image segmentation, are computationally intensive. Implementing and training these models may require large computational resources. This challenge necessitates the use of optimized architectures, model compression techniques, or cloud-based solutions to improve efficiency without compromising accuracy.

4.8 Generalisation to Different Patient Populations

Models trained on data from one population may not generalise well to patients from different demographics or with different health conditions. Ensuring the robustness and generalisability of segmentation models is an ongoing challenge.

5. DATA PREPROCEESING

5.1 Data Collection

Medical image data is acquired for various purposes, such as diagnosis, therapy planning, intraoperative navigation, postoperative monitoring, and biomedical research [53]. The use of medical datasets needs to go through a rigorous clearance process such as obtaining ethical clearance [54]. Data ethics and medical image analysis have become an inseparable entity as medical data is based on data sets provided by medical centres or research laboratories. The processes they employ to collect, store, and share data for medical data must take into account the privacy of patient data [55].

Data collection is the process of systematically and reproducibly collecting and measuring variables to answer research questions, test hypotheses, or evaluate results [56]. Obtaining data is the most important step in a research study [56]. Medical image and video datasets can support biomedical research through training machine learning algorithms, particularly through image recognition and classification. This can be applied to problems in digital health informatics, such as disease detection, diagnosis and screening.

5.2 Data Annotation

The annotation process aims to transfer human knowledge to artificial intelligence models by summarising and assigning pre-defined labels to digital data content [56]. Annotation tools are characterised by the tasks covered, the functionality provided, and the features supported such as pre-processing and automatic labelling. These tools allow users to label objects of interest in frames by supporting three modalities: manual, semi-automatic, and automatic.

Choosing the right tool is important as it greatly affects the quality of the data and the time taken to complete it. Image annotation is a key element of medical image-based computer-aided diagnosis systems [57]. Annotation refers to the process of manually defining regions in an image and then adding metaphors or comments of those specific regions in the form of text. Image annotation allows the assignment of metadata to digital images as captions or keywords. This reduces the latent information in the scan into clearer details, enabling accurate clinical decision-making. Figure 1 illustrates the annotation process for Diabetic Foot Ulcers, where lesion regions are manually or semi-automatically labeled to create ground truth masks for segmentation models [58].

(a) Images (b) Ground Truth

Fig. 1. Diabetic Foot Ulcer Annotation

Image annotation plays a crucial role in the development of machine learning models for Diabetic Foot Ulcer segmentation, as it directly impacts the accuracy and effectiveness of the model's predictions. Annotation involves the precise tagging and labeling of ulcer regions in medical images, creating a ground truth dataset that serves as a benchmark for training and evaluating segmentation algorithms. This process can be performed manually by expert clinicians or semi-automatically using annotation tools, ensuring that the model learns to differentiate ulcerated areas from healthy tissue. High-quality annotations are essential for reducing false positives and false negatives, thereby improving the model's ability to generalize across different patient cases. Additionally, consistent and well-annotated datasets contribute to more reliable model performance, making them fundamental in medical AI applications.

5.3 Data Splitting

Separation of data into training set, testing set, and validation set is an important step in the process of training and evaluating machine learning models [59]. This separation helps to objectively measure model performance and prevent overfitting. The training set is used to train the model on data with known labels with a composition ranging from 70- 80% of the total data. The validation set is used to adjust the model parameters and select the best model. The validation set serves as an intermediate evaluation during training to avoid overfitting. The composition of the validation set ranges from 10-20% of the total data.Test set is used to measure the performance of the trained and adjusted model. Test data should not be used in the process of training or adjusting the model. Independently, the test set is used to evaluate the ability of the model to make predictions on data that has never been seen before.

5.4 Data Augmentation

Data augmentation is a popular technique that helps improve the generalisation ability of deep neural networks, and can be considered as implicit regularization [60]. Traditionally, data augmentation approaches have been applied to increase the size of training sets, so that large-capacity learners can benefit from more representative training data. Data augmentation is the process of modifying or manipulating an image, so that the original image in standard form will be changed in shape and position. Data augmentation aims to enable machines to learn and recognise from a variety of different images and can also be used to augment data. In most cases, the use of data augmentation improves the performance of the model. The improvement occurs because the machine manages to recognise more objects of various shapes and patterns. As can be seen in the image below, a single image can produce many new images with various positions. In the affine approach, the existing image data undergoes different operations (rotation, zoom, cropping, flipping, or translation) to increase the number of training examples. Figure 2 illustrates the process of image augmentation, which is used to enhance the diversity of training data and improve model generalization [61].

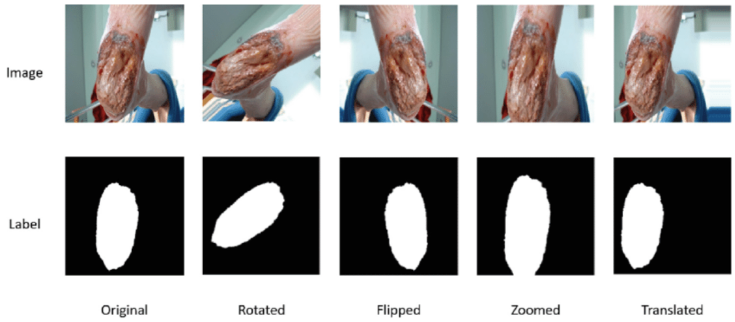

Fig. 2. Image Augmentation

5.5 Data Resizing

Resizing an image is one of the most fundamental operations on an image [62]. There are several approaches to doing so in order to maximise the quality of the resulting image. Image resizing is an important pre-processing step in computer vision. In principle, deep learning models train faster on small images. Larger input images require the neural network to learn from four times as many pixels, and this increases the training time for the architecture [63].

5.6 Data Normalization

Image normalisation ensures optimal comparison across data acquisition methods and texture examples [64]. Normalisation of pixel values (intensity) is recommended for imaging modalities that do not correspond to absolute physical quantities. Various advanced strategies have been proposed to normalise values and are often modality specific. Histogram equalisation, also known as image normalisation, can be used to make contrast adjustments to the image so that the image intensity values span the entire intensity range [65]. In addition, this larger range of intensity values allows for a sharper distinction between dark and light regions.

5.7 Handling Class Imbalance

Unbalanced data distributions naturally arise in many applications where positive classes occur with reduced frequency, including data found in disease diagnosis, fraud detection, computer security, and image recognition [66]. The imbalance problem was first recognised almost three decades ago and is still an important challenge for at least traditional learning approaches [67]. The imbalanced learning problem is related to the performance of learning algorithms in the presence of asymmetric class distributions [68]. Due to the complex characteristics of imbalanced datasets, learning from such data requires new algorithms and understanding to efficiently transform a large amount of initial data into a suitable dataset.

6. POTENTIAL TECHNOLOGY USED

6.1 U-Net

U-Net is an image segmentation technique developed primarily for image segmentation tasks [69]. These features give U-Net a high usability in the medical imaging community and resulted in the widespread adoption of U-Net as the primary tool for segmentation tasks in medical imaging [70]. The success of U-Net is evident in its widespread use in almost all major image modalities, ranging from CT scans and MRIs to X-rays and microscopy.

Semantic segmentation is the classification of features in images based on pixels. Due to the lack of image details, it is impossible to obtain precise boundaries using the semantic feature information of the image. U-Net comes up with the concept of skip connection to combine the feature maps of low resolution and high-resolution images [71]. U-Net is one of the CNN architectures that has a simple encoder and decoder network that forms the letter U. The U-Net model can work on small datasets. Figure 3 is a representation of the U-Net architecture [71].

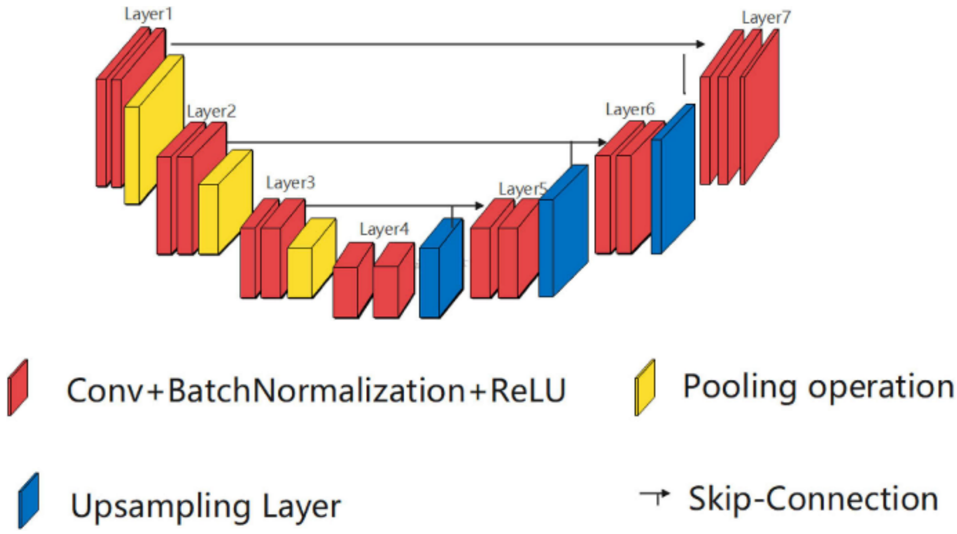

Fig 3. U-Net Architecture

Figure 3 illustrates the U-shaped U-Net architecture, which consists of three key components: the contracting path, the expansive path, and skip connections. The contracting path is responsible for capturing high-level spatial features through a series of convolutional and downsampling operations, progressively reducing the spatial dimensions while increasing the depth of feature representations. The expansive path, on the other hand, aims to reconstruct the segmented output by applying upsampling operations to recover spatial information. To preserve fine-grained details and improve localization accuracy, skip connections are employed to directly transfer features from corresponding layers in the contracting path to the expansive path. These skip connections help mitigate information loss during downsampling and enhance the model's ability to precisely identify the regions of interest (ROIs) within the input images.

6.1.1 Encoder

The U-Net architecture has a symmetrical encoder-decoder structure, with the encoder part responsible for extracting features from the input image. The encoder part starts with a series of convolution layers that aim to extract features from the input image. Each convolution layer is followed by an activation function (usually ReLU) to introduce non-linearity into the model. A normalisation layer, such as Batch Normalisation, is often used after the convolution layer to help stabilise the training and speed up convergence. To extract features hierarchically and reduce spatial resolution, downsampling operations such as max pooling or strided convolution are used. Downsampling helps the model to understand the image structure more generally. Each layer in the encoder section is directly connected to the corresponding layer in the decoder section through skip connections. Skip connections help retain high spatial information during the decoding process and help overcome the problem of loss of detail during downsampling. An encoder block consists of one or more convolution layers, followed by an activation function and a normalisation layer. The number and size of convolution layers in each block may vary depending on the complexity of the task and the nature of the dataset. In general, the encoder part of U-Net has a multilevel architecture that extracts features hierarchically from the input image. Each encoder block reduces the spatial dimension and increases the number of features extracted, allowing the model to understand image structures and patterns with varying degrees of complexity. The skip connection scheme allows high spatial information to be reintegrated into the decoding process, helping to produce high-detail segmentation in the output image.

6.1.2 Bottleneck

The bottleneck in U-Net is the part where the number of features peaks after downsampling and before upsampling. In this part, the downsampled image is converted into a more compact representation, and then upsampled again to produce the segmented output image. This bottleneck is at the centre of the U-Net architecture and plays a key role in bringing together information from both sides (encoder and decoder). The bottleneck part of the U-Net aims to significantly reduce the image dimension. This is usually achieved by using convolution layers with larger strides or pooling operations that cause downsampling. Although the spatial dimension has been reduced, the number of features represented in the bottleneck remains high. This high feature representation helps in storing important information about the image that is required for reconstruction at the decoder. The skip connection path from encoder to decoder inserts high spatial information into the bottleneck section. This helps the model maintain high detail and fixes the problem of information loss during the downsampling process. At the bottleneck, there are often multiple convolution layers that are responsible for extracting increasingly abstract representations of the image. Convolution layers here can help in describing more complex features. The number of channels (features) at the bottleneck usually increases to allow the model to store rich information while maintaining a lower spatial dimension.

6.1.3 Decoder

The decoder part of the U-Net architecture is responsible for evolving the feature representation that has been extracted by the encoder and rebuilding the segmented output image. The decoder performs an upsampling operation to restore the spatial dimension of the image and produce an output that has the same resolution as the input image. The upsampling operation is performed to restore the spatial dimension of the image after the downsampling process at the encoder. Upsampling can be done using techniques such as bilinear interpolation or transposed convolution (transpose convolution or deconvolution). Each layer in the decoder section is connected to the corresponding layer in the encoder section via a skip connection. The information from the encoder, which has been inserted during the downsampling process, is concatenated with the features generated by the decoder to aid the reconstruction of the high-detail image. The decoder section generally consists of a series of convolution layers that aim to develop increasingly rich and complex feature representations. These convolution layers help in describing the features required for the segmentation task. Each convolution layer is followed by an activation function, such as ReLU, to introduce non-linearity into the model. A normalisation layer, such as Batch Normalisation, can be used to aid in training stabilisation and convergence acceleration. The output layer in the decoder section produces a segmented output image. Typically, this output has the number of channels corresponding to the number of classes to be segmented. Overall, the decoder part of U-Net works to resize the image from the higher feature representation in the encoder part. The use of skip connection paths allows high spatial information to be re-included during the upsampling process, helping in maintaining the detail and quality of the segmentation.

6.1.4 Skip Connections

The skip connection in U-Net is the key element that distinguishes this architecture from other image segmentation architectures. Skip connections, also known as shortcut paths, feed information from the encoder section directly to the decoder section. This makes it possible to retain high spatial information during the decoding process and helps overcome the problem of loss of detail during downsampling operations in the encoder section. each box represents a layer or operation in the U-Net architecture. The skip connection path is represented by a line connecting the corresponding layers between the encoder and decoder sections. Information from the encoder section (high feature representation) is reinserted into the decoder section to aid in the reconstruction of the high-detail image during the upsampling process.

6.1.5 Final Layer

In the U-Net architecture, the output layer or "final layer" is the last part of the network that produces the prediction or segmentation results. In semantic segmentation tasks, the softmax activation function is commonly used in the output layer. Softmax generates a probability distribution for each possible class or category for each pixel in the image. This allows each pixel to be classified as belonging to a particular class. The number of output channels in the output layer usually corresponds to the number of classes or categories that the model should identify. The output layer is also associated with the loss function used during model training. This loss function compares the model predictions with the ground truth labels to measure the extent to which the model matches the desired target.

6.2 Varian U-Net

Since U-Net was first introduced in 2015 [72], many variants of U-Net have been developed to meet specific needs and challenges in medical image segmentation tasks and other fields. Some of these variants introduce changes to the basic architecture of U-Net to improve the performance and robustness of the model. Here are some of the U-Net variants that have significant performance.

6.2.1 U-Net++ (Nested U-Net)

U-Net++ is a modification of the U-Net architecture that has been introduced to improve performance in image segmentation tasks [73]. U-Net itself has been widely used in various fields, especially in the field of medical image processing. U-Net++ introduces a more complex skip connection scheme. It includes longer and more numerous skip pathways, which is expected to help the model capture more complex and abstract features. Hyperconnection is a concept in U-Net++ that connects each level to another, creating a more information-rich network. U-Net++ uses dilated convolutions to capture global context without sacrificing local resolution. This helps the model to understand the broader context without losing important local details. U-Net++ utilises residual connections, which is a concept introduced by ResNet. Residual connections help prevent training problems that in some cases can occur in very deep networks. UNet++ is designed to improve U-Net's ability to capture more complex features and context, which can be especially useful in challenging image segmentation tasks. While this concept can provide improved performance, it is important to remember that the success of the model also depends on the quality and size of the training dataset, as well as the right training parameters.

6.2.2 Residual U-Net (ResU-Net)

Residual U-Net is a variant of the U-Net architecture that utilises the concept of residual connections [74]. Residual connections were first introduced in the ResNet (Residual Networks) architecture to address the performance degradation problem that arises when the model becomes deeper [75]. Residual U-Net utilises a block of residual connections. Residual connections allow information to "jump" or "flow back" through the network layers, enabling faster learning and helping to prevent a training problem called performance degradation. Residual U-Net has an encoder-decoder structure. The encoder part serves to extract features from the image, while the decoder part is used to high-resolve the prediction results. Residual U-Net also retains skip connections that allow information from the encoder layer to be directly connected to the decoder layer. This helps the model to incorporate contextual information from different resolution levels. Some Residual U-Net implementations also utilise dilated convolutional blocks to increase the network's capacity to capture spatial context. Residual U-Net is generally designed to enhance the deep learning capability of the network by using residual connections. Residual connections help in learning better representations, especially as the network becomes deeper, and this can be beneficial in image segmentation tasks where understanding complex context and features is required.

6.2.3 Attention U-Net

Attention U-Net is a variant of the U-Net architecture that incorporates an attention mechanism element [76]. The attention mechanism allows the model to "focus" on specific parts of the input during the learning process, allowing the model to give different weights to different features [77]. Attention mechanisms are implemented to give different levels of attention to different parts of the input. This allows the model to pay more attention to relevant features and ignore parts that may not be relevant for a particular task. Attention U-Net maintains the general structure of U-Net, including encoder-decoder architecture and skip connections. Some implementations of Attention U-Net may use self-attention or spatial attention. Self-attention gives the model the ability to adjust the weights to its own input features, while spatial attention can give different weights to certain spatial parts of the input. Some implementations of Attention U-Net may also include blocks of residual connections to aid in learning better representations, especially as the network becomes deeper. Attention U-Net is designed to improve the model's ability to capture important relationships and information in images, especially in image segmentation tasks. By incorporating an attention mechanism, the model can focus on important details which can improve performance in more complex segmentation tasks or in conditions where there is significant distraction or noise.

6.2.4 Recurrent U-Net (ReU-Net)

Recurrent U-Net is a variant of the U-Net architecture that incorporates recurrent elements [78]. This is done to allow the model to integrate contextual information from previous iterations, which can be beneficial in image segmentation tasks or temporal context understanding. Recurrent U-Nets include recurrent layers or blocks, such as LSTM (Long ShortTerm Memory) cells or GRU (Gated Recurrent Unit) cells, at certain levels in the architecture. These recurrent layers allow the model to store and access information from previous iterations, enabling temporal contextual understanding. Like U-Net, Recurrent U-Net maintains the general structure of an encoder-decoder. The encoder part is for extracting features and the decoder part is for high-resolution prediction results. Recurrent U-Net can also utilise skip connections to combine information from different levels of resolution. By incorporating recurrent, Recurrent U-Net can combine spatial information from the U-Net with temporal information from previous iterations. Recurrent U-Net can be useful in contexts where temporal information plays an important role, such as in video segmentation tasks or other tasks that involve understanding temporal context. However, the implementation and success of the model depends on the specific properties of the data and the task at hand.

6.2.5 V-Net

V-Net (Volumetric Convolutional Neural Network) is a convolutional neural network architecture specialised for volumetric segmentation tasks, especially in the context of medical images, such as 3D medical image processing from CT (Computed Tomography) or MRI (Magnetic Resonance Imaging) [79]. V-Net was developed specifically to handle volume information reflected in medical data. V-Net uses a 3D convolutional layer to handle volumetric data. 3D convolution allows the model to directly extract features from the spatial volume. V-Net has an encoder-decoder structure. The encoder part is for extracting features and the decoder part is for high-resolution prediction results. V-Net also utilises skip connections to combine information from different resolution levels. This helps to correct the problem of spatial information loss during the down-sampling process. V-Net typically uses batch normalisation followed by a ReLU (Rectified Linear Unit) activation function for each convolution layer. This helps in accelerating training and preventing vanishing gradient problems. The output layer uses softmax activation to assign segmentation probabilities to each voxel in the volume. V-Net was developed specifically to handle volumetric data in the context of medical images. Its main advantage is its ability to handle the volume structure of the human body and produce accurate segmentation in tasks such as segmentation of specific organs or anatomical structures in 3D medical images.

6.2.6 LinkNet

LinkNet is a convolutional neural network architecture developed specifically for image segmentation task [80]. Created by a company called 'Adequate Intelligence' (now known as Insight Robotics), LinkNet is designed to provide accurate and efficient segmentation results with relatively low computational requirements. LinkNet has an encoder-decoder structure similar to the U-Net architecture [81]. The encoder part is responsible for extracting features from the image, while the decoder part is used for high-resolution prediction results. LinkNet uses skip connections that connect the same level of resolution between the encoder and decoder. This helps in overcoming the problem of spatial information loss during the down-sampling process. LinkNet uses a bottleneck block consisting of separable convolution layers. This block helps to reduce the number of parameters and computation required, thus improving efficiency. As in many convolutional neural network architectures, LinkNet uses batch normalisation followed by a ReLU activation function for each convolution layer. A unique feature of LinkNet is the use of modules called "Link Modules." These modules are used to link different resolution levels between the encoder and decoder. Link Modules help in transferring relevant information from the encoder to the decoder. The upsampling process at the decoder is done using skip connections from the encoder, which helps in incorporating contextual information from lower resolution levels. LinkNet is generally seen as an efficient and effective architecture for image segmentation tasks, especially when computational requirements are an important consideration. LinkNet's advantage lies in the balance between accuracy and computational efficiency, making it suitable for a wide range of applications, including in the fields of computer vision and image processing.

6.2.7 DoubleU-Net

DoubleU-Net is a modification of the U-Net architecture developed to improve segmentation capabilities in medical tasks, specifically on 2D or 3D medical data such as CT or MRI images [82]. The name "DoubleU" may refer to the doubling or repetition of some elements in the architecture. DoubleU-Net may double some parts of the U-Net structure, such as U-Net Blocks [82]. This may include blocks in both the encoder and decoder, increasing the capacity of the model to capture more complex features. Structures with two U-Net Blocks can generate more complex skip connections, allowing better information exchange between different resolution levels. DoubleU-Net may retain the use of batch normalisation and activation functions (such as ReLU) at each convolution layer, in accordance with common practice in convolutional neural network architectures. Regularization techniques such as dropout or normalisation techniques such as batch normalisation layers are used to improve generalisation and training stability. The design of DoubleU-Net can be customised based on the specific task and characteristics of the medical dataset used. DoubleU-Net was designed with the aim of improving the model's ability to handle segmentation tasks, especially on medical data. The effectiveness of this model will largely depend on the fit with the task problem and the characteristics of the data at hand. Therefore, experimentation and evaluation on specific relevant datasets will provide further insights into the performance of DoubleU-Net in specific contexts

6.3 Related Research on DFU Image Segmentation

Several previous studies have been conducted by researchers in the implementation of deep learning in wound segmentation including DFU segmentation. There are several segmentation methods that have been carried out and have advantages and disadvantages of research. Potential method development from the weaknesses found can be applied in future research.

Research conducted by Goyal (2019) [17] found a research problem, namely the need for DFU segmentation technology due to the difficulty in detecting and localising diabetic foot ulcers in foot images due to inter-class similarities and intraclass variations in terms of colour, size, shape, texture, and location between different DFU classes. Another problem is the limitation of research data and the scalability of the DFU diagnosis system to various devices, platforms, and operating systems, to ensure wide accessibility and usability. The method used in this research is to use convolutional neural networks (CNN) to localise DFU in real time. Deep learning models were trained using different object localisation metaarchitectures, including Faster R-CNN and Inception V2, and evaluated using five-fold cross validation. The results stated that the overall Faster R-CNN with Inception V2 model achieved an average mean precision of 91.8% in five-fold crossvalidation. The models were also tested in real-time on NVIDIA Jetson TX2 and smart phone applications, which showed efficient predictions. Although this study has advantages in performance, there are still weaknesses in the form of limited standard data, integration of additional diagnostic features such as the combination of image features with diagnostic features, such as patient ethnicity, presence of ischaemia, DFU depth in tendons, and neuropathy and the system does not yet meet scalability across devices and platforms.

Research that has been conducted by Zhao (2019) [35] found a research problem that is how visually similar wounds are difficult to classify accurately. The method used in this study is a detailed classification approach, specifically the Bilinear Convolutional Neural Network (Bi-CNN) architecture, to assess the depth and amount of granulation tissue of diabetic wounds. The results showed that the proposed Bilinear Convolutional Neural Network (Bi-CNN) approach outperformed other CNN architectures in assessing diabetic wounds in terms of accuracy. The accuracy of assessing wound depth and the amount of granulation tissue was 84.6%. Although the accuracy value increased compared to previous studies, there were still some misclassifications due to factors such as unstable lighting, blurring, low resolution, and controversial labels

Research conducted by Rania (2020) [18] tries to create an automatic DFU segmentation method with the U-Net network to overcome accuracy errors when using manual segmentation. The U-Net method is compared to other deep learning networks such as V-net and Seg-Net. The results show that the U-Net network has the best performance with a dice coefficient of 97.25% and an IoU of 94.86%. Despite having the advantage of getting good performance values, this research still involves a small dataset thus limiting the accuracy and generalisation of the results.

Research that has been conducted by Wang (2020) [30] with the research problem is the development of automatic segmentation of wound areas in natural images. Another problem found is that deep learning methods require large data sets for training, and there is currently no public data set large enough to train deep learning models for wound segmentation. The method used in this document is an automated wound segmentation framework based on MobileNetsV2 and connected component labelling. The research also compared their proposed framework based on MobileNetsV2 with other popular deep learning models such as FCN-VGG-16, SegNet, U-Net, and Mask-RCNN. The results show that their proposed framework achieves high accuracy with an average Dice score of 90.47% and its performance is comparable to other deeper neural networks. Despite the good performance of the algorithm, this research still has the potential to improve its performance value.

Research conducted by Heras (2022) [20] saw the problem of observer perception variation in manual segmentation. This motivates researchers to create automatic segmentation with computer-aided algorithms. The methods used to create automatic segmentation utilise hybrid random classifiers, decision trees, gaussian naive bayes, logistic regression and SVM techniques. The proposed algorithm consists of two stages: ulcer region segmentation and post-processing of segmentation results. The algorithm achieved a Jaccard Index of 0.81, accuracy of 0.94, recall of 0.86, precision of 0.91, and F1 score of 0.88. Despite the promising performance of the algorithm, this study still has the disadvantage of lacking information on the size and diversity of the datasets used for training and evaluation, which may affect the generalisability of the algorithm's performance. Another weakness is that the comparison of learning models is limited, with only five models compared in this study. Other advanced models may achieve better results.

Research conducted by Huang (2022) [36] tries to make DFU segmentation technology more accurate and systematic. The method used in the study is a combination of deep neural network technology, convolutional neural network (CNN), object recognition, transfer learning, and image analysis technology to analyse and classify diabetic foot wounds. The results showed that the proposed image segmentation technology for diabetic foot ulcers (DFUs) achieved a high accuracy rate of 90% in assessing wound image detection. The study compared different deep neural network models and found that ResNet showed the highest learning accuracy. In addition, the study used GrabCut and SURF algorithms to identify wounds and calculate various characteristics of each wound, such as area and perimeter. Although the researchers claimed good accuracy, the results and effectiveness of this study were not widely discussed or validated, making it difficult to determine the overall reliability and generalisability of the proposed technology. Further research and validation may be required.

Research conducted by Liu (2022) [83] attempts to address the problem of accurately diagnosing infection and ischaemia in diabetic foot ulcers (DFUs) using a deep learning model. Other challenges include high similarity and variation in infection and ischaemia classes, non-standardised imaging conditions, and lack of patient demographic information in the dataset. This study utilises the EfficientNets neural network architecture to develop a deep learning model for diagnosing infection and ischaemia in diabetic foot ulcers. Research on the diagnosis of infection and ischaemia in diabetic foot ulcers using the EfficientNet model achieved excellent results. The EfficientNet model achieved 99% accuracy in ischaemia classification and 98% accuracy in infection classification. These results outperformed other basic models, including ResNet and Inception, as well as the previous state-of-the-art CNN Ensemble model. In addition, the EfficientNet model was significantly faster than the baseline models, and took only a very short time to classify the test images. Despite the high performance in classification, the model may lack interpretability. This is due to the difficulty in understanding the specific features or factors that influence the model's classification decision. This can make it difficult to explain the reasoning behind the model's predictions. The performance of the EfficientNet model is highly dependent on the availability of sufficient data. If a model is trained on a specific data set, it may not generalise well to different data sets or real-world scenarios. This limits its applicability in various situations.

Research has been conducted by Ahsan (2023) [32] who tried to create a more reliable and precise segmentation technology in segmentation and classification of DFU. This research was also constrained by the scarcity of data and labelling by medical experts, as well as the asymmetrical and multicoloured shape of skin lesions in DFUs, posing additional challenges. A CNN-based deep learning architecture was used in this study, including ResNet50, to achieve accurate classification of DFUs. Transfer learning with a fine-tune approach is leveraged to overcome data scarcity and labelling by medical experts in the field of medical imaging. Data augmentation techniques, such as rotation, flipping, scaling, translation, mirroring, and shearing, were applied to increase the input size of the target domain. CNN models were trained and refined using the DFU2020 dataset to classify ischaemia and infection in DFUs. This study used multiple CNN architectures and adjusted weights to classify DFUs into ischaemia and infection categories. The ResNet50 architecture achieved the highest accuracy of 99.49% for ischaemia classification and 84.76% for infection classification. Performance comparison with other techniques showed that the proposed strategy outperformed previous studies in terms of accuracy, F1-Score, MCC, and AUC. Despite having good performance results, this study still found some drawbacks such as data scarcity and labelling, high computational resources when performing training, and the deep learning model is highly dependent on the quality and representativeness of the training data.

Research conducted by Kucharski (2023) [84] proposed an end-to-end ensemble fully convolutional network (DFU-Ens). The model consists of 3 main modules namely the U-Net module (DFU-Seg), a hybrid approach (DFU- Det) containing YOLOv4 based detection module and DETR Vision Transformer detection approach. The ensemble solution is based on a high-ranking strategy that combines DFU-Seg with a hybrid solution of bounding box detection and patch segmentation. On the DFUC2022 validation set, the study achieved Dice scores of 0.643 for the ensemble approach, 0.648 for DFUSeg, and 0.556 and 0.581 for the hybrid approach based on YOLOv4 and DETR respectively, with DETR having the best sensitivity. On a later published test set, DFU-Seg achieved a Dice score of 0.67 while the ensemble method achieved 0.66.

Research conducted by Shah (2023) [85] proposed CADFU (Computer-Aided Diagnosis System for Foot Ulcers) a pioneering diabetic foot ulcer diagnosis system. This method aims to detect and categorise ulcers and similar chronic wounds in medical images. The built model consists of two different algorithms namely DhuNeT which is an innovative Two-Phase Hyperactive U-Net architecture for the segmentation task and YOLOv8 for the wound detection task. DHuNeT is a combination of stacking two UNet autoencoders sequentially. Hyperactive transmission of information from the first UNet to the second UNet is the main idea of DHuNeT. The first UNet provides the second UNet with the features it has learnt, and both UNets combine their learnt features to create new, more accurate, and effective features. The results of this study achieved good performance measures, especially in terms of dice coefficient and precision, with segmentation scores of 85% and 92.6%, respectively

Based on previous research, it appears that the U-Net architecture is still popularly used for medical image segmentation. The researcher explained that there is still potential to improve the accuracy performance of DFU segmentation models. There is also the problem of limited datasets, especially DFU images. Both of these are challenges for future research. Table 3 represents the results of previous research on DFU image segmentation.

6.4 Potential Fuzzy Logic and U-Net in Image Segmentation

U-Net is one of the CNN architectures that uses encoder-decoder networks and has achieved successful performance in the field of medical image segmentation [86]. The architecture owned by U-Net is certainly not much different from that of CNN. U-Net consists of Convolution Operation, Max Pooling, ReLU Activation, Concatenation and Up Sampling Layers and three sections such as contraction, bottleneck, and expansion section [87].

The potential of fuzzy logic can certainly be applied to the U-Net architecture such as the application of fuzzy logic to CNN. Medical images are generally noised [44] and fuzzy logic can be used as a solution to this problem. Research conducted by [88] tries to remove Gaussian-impulsive mixed noise in medical images using fuzzy logic-based parallel filters. This research produced an efficient parallel algorithm based on fuzzy logic to remove Gaussian impulsive noise in CT medical images. This algorithm shows a reduction in computation time so that the proposed technique can be applied for real-time medical image filtering.

The encoder in U-Net is used to perform feature extraction [89]. The encoder's task is to represent a low-dimensional input image that contains only the most important information of an image. The encoder part of U-Net can be enhanced by utilising fuzzy logic to customise the feature extraction process. For example, the output of the encoder layer can be processed with fuzzy logic to determine the relative weights of various features.

After the feature extraction process with the encoder, the image has low dimensions so a decoder process is needed. The decoder serves to reconstruct images from feature extraction with increased dimensions [90]. In the decoder part, fuzzy logic can be used to adapt the results of the skip connection path and combine them with the decoding results at each resolution level. This can help in adaptive adjustment to different levels of complexity in different parts of the image.

The skip connection path is a key feature of U-Net. During the decoding process, it allows spatial information lost during downsampling operations at the encoder to be reintegrated into the higher representation. This helps to overcome the problem of loss of detail during the decoding process [91]. Fuzzy logic can be applied to skip connections to determine the extent to which information from the encoder layer should be used in the decoding process. This can help address uncertainty in the image and enable better adaptation to structural variations.

Opportunities that can still be applied with fuzzy logic include fuzzy-based parameter adjustment. Parameters in U-Net, such as weights and biases, can be adjusted using fuzzy logic. This allows the model to dynamically adjust the parameters based on complex medical image conditions. Fuzzy logic also has an opportunity for information fusion. Fuzzy logic can be used to combine information from different levels of resolution, not only spatially but also in terms of information quality and relevance. This can increase the capacity of the model to understand both global and local contexts.

Another application of fuzzy logic in the U-Net architecture is the fuzzy pooling layer. This is inspired by the research done by [92] with fuzzy-pooling method that successfully found better representative features for ultrasound image classification. The designed fuzzy-pooling based CNN architecture works on enhanced images to extract representative features for classification. The feature map taken by the pooling layer as input is the output of the convolution layer. The convolution layer performs a convolution operation, then the result passes through an activation function, specifically the Rectified Linear Unit (ReLU). ReLU always gives non-negative values as output, which formulates the pooling layer's input feature map as a collection of non-negative values. The conversion of non-negative values from feature maps into fuzzy values is initiated through fuzzy membership functions. This research performs fuzzification on the feature map and obtains three versions of the fuzzified feature map. The defuzzification stage reduces the dimension of the patch. This research has also used different defuzzification methods and found Centre of Gravity (COG) to be an effective technique. The Xception model achieved the best classification results with 97.2% accuracy.

Table 3. Summary of Research Related to DFU Image Segmentation

Work	Dataset	Method	Result
Goyal (2019) [17] Robust Methods for Real-Time Diabetic Foot Ulcer Detection and Localization on Mobile Devices	Lancashire Teaching Hospitals Dataset	CNN, Faster R-CNN dan Inception V2	Precission 91.8%
Zhao (2019) [35] Fine-Grained Diabetic Wound Depth and Granulation Tissue Amount Assessment Using Bilinear Convolutional Neural Network	University of Massachusetts Medical School DFU Dataset	Bilinear Convolutional Neural Network	Accuracy 84,6%
Rania (2020) [18] Semantic segmentation of diabetic foot ulcer images : Dealing with small dataset in dl approaches	Primary Dataset	U-Net, V-Net and Seg-Net	Dice Coefficient 97,25% and IoU 94,86%
Wang (2020) [30] Fully automatic wound segmentation with deep convolutional neural networks,	Advancing the Zenith of Healthcare (AZH) Wound and Vascular Center, Milwaukee, WI Dataset	MobileNetsV2, FCN-VGG-16, SegNet, U-Net, dan Mask-RCNN	Dice Coefficient 90,47%
Heras (2022) [20] Diabetic foot ulcer segmentation using logistic regression, DBSCAN clustering and morphological operators	Own annotated dataset from images of Cuban patients	Hybrid random classifiers, decision trees, gaussian naive bayes, logistic regression and SVM	Jaccard Index of 0.81, accuracy of 0.94, recall of 0.86, precision of 0.91, and F1 score of 0.88
Huang (2022) [36] Image segmentation using transfer learning and Fast R-CNN for diabetic foot wound treatments	International Working Group on the Diabetic Foot Dataset	Fast R-CNN	Accuracy 90%
Liu (2022) [83] Diabetic Foot Ulcer Ischemia and Infection Classification Using EfficientNet Deep Learning Models	Diabetic Foot Ulcers Grand Challenge (DFUC) 2021	EfficientNets	99% accuracy in ischaemia classification and 98% accuracy in infection classification
Ahsan (2023) [32] A Deep Learning Approach for Diabetic Foot Ulcer Classification and Recognition	Diabetic Foot Ulcer 2020 (DFU2020) dataset	CNN, ResNet50	The highest accuracy was 99.49% for ischaemia classification and 84.76% for infection classification.
Kucharski (2023) [84] DFU-Ens: End-to-End Diabetic Foot Ulcer Segmentation Framework with Vision Transformer Based Detection BT - Diabetic Foot Ulcers Grand Challenge	Diabetic Foot Ulcer 2022 (DFU2022) dataset	U-Net (DFU-Seg), hybrid approach (DFU-Det) containing YOLOv4	Dice scores of 0.643 for the ensemble approach, 0.648 for DFU-Seg, and 0.556 and 0.581 for the hybrid approach
Shah (2023) [85] CADFU for Dermatologists: A Novel Chronic Wounds & Ulcers Diagnosis System with DHuNeT (Dual-Phase Hyperactive UNet) and YOLOv8 Algorithm	MICCAI 2021 Foot Ulcer Segmentation (FUSeg) Challenge	U-Net Hiperaktif Dua Fase and YOLOv8	Dice Coefficient 85% and Precission 92.6%

Equally interesting is the potential use of fuzzy logic on the U-Net activation function. This was inspired by a study [24] that proposed to combine deep learning for feature extraction and single-layered RVFL for classification. In addition, this study highlights the use of fuzzy activation functions to handle outliers in MRI images. Fuzzy activation function (FAF) is used in the hidden layer of the deep random vector functional link (DRVFL). FAF is used to transform the features extracted from MRI images into a non-linear space and remove outliers present in the data. FAF works by mapping the input features to a set of membership values, which allows for a more flexible and adaptive representation of the data. This fuzzy mapping helps capture uncertainty and imprecision in the input features, which is particularly useful in dealing with outliers and noise that may be present in MRI images. The proposed classification model can effectively handle the variation and complexity of the input data, thus improving the accuracy in diagnosing Alzheimer's disease at an early stage.

The FAF-based DRVFL network provides a robust and efficient framework for automated diagnosis, assisting clinicians in making accurate and timely decisions. The performance of the proposed model is compared with other noniterative classifiers and activation functions. The results show that the proposed model achieves high accuracy in classifying different classes of patients based on MRI structural images. The proposed s-FAF activation function outperforms other activation functions in terms of accuracy.

7. STATE OF THE ART

Based on various previous studies on the use of artificial intelligence in the field of diabetic foot wound image segmentation, various research problems were found, including the limited availability of training data and the need for further clinical trials to make the segmentation process more accurate. This data limitation is stated by [17] [30] [31] [19] [32]. The data used by various previous studies is also still limited to the challenges of data heterogeneity such as image quality [83], variations in the shape of skin lesions [32] and the subjectivity of annotating diabetic foot wound images from a variety of observers [20].

Another problem that still needs to be found is the problem of different segmentation evaluations from various researchers. There are researchers who produce high accuracy values such as those conducted by [18] where the F1-Score results reach a value of 97.25%, IoU 94.86% and accuracy 94.96%. After further study, the results of the high evaluation value are due to the very small number of datasets, which is below 100 so it is still considered not accurate enough. There are also high evaluation results such as those conducted by [20] with a Jaccard Index value of 0.81, accuracy of 0.94, recall of 0.86, precision of 0.91, and F1 score of 0.88. This also still needs to be reviewed because the research does not explain in more detail regarding the dataset used. There are also several studies that can still be improved in terms of accuracy such as research [35] which still has an accuracy value of 84.6%, with a dice coefficient value of 74.1%, [93] with a dice coefficient value of 0.648.

The difference in accuracy, *f1-score*, *precission*, *jaccard index*, *IoU* and so on as a model evaluation, is overcome by developing a new model that utilises fuzzy logic and qmetaheuristic optimisation. Fuzzy logic has been applied in many previous studies to optimise the performance of deep learning. The utilisation of fuzzy logic is proposed because there are many unclear data or processes carried out by deep learning [94]. Segmentation fuzziness such as reading specific *pixel* areas to make it easier to read the boundary of the image also has its own challenges [95]. There has also been a lot of research using fuzzy on image segmentation to detect image edges as done by [96] [22] There are also studies that try to optimise the *hidden layer* in image handling such as fuzzy logic applied to the pooling layer [97] or convolutional layer [98].

8. PERFORMANCE ANALYSIS

Evaluation of semantic segmentation models is an important process to assess the performance and accuracy of the model in determining the boundaries and classes of objects in an image. The following are some evaluation metrics that are commonly used to evaluate semantic segmentation models.

8.1 Intersection over Union (IoU)

The intersection over union (IoU) score, also called Jaccard Index, is one of the most fundamental evaluation methods in machine learning. The original IoU computation cannot provide a non-zero gradient so it cannot be optimised directly with current deep learning methods [99]. IoU is the overlap area between the predicted segmentation and the ground truth divided by the union area between the predicted segmentation and the ground truth [100]. The IoU formula can be expressed as follows.

$$IoU = \frac{Intersection}{Union} \quad (1)$$

Intersection is the area where the prediction result and ground truth overlap, and Union is the total area of the two regions. IoU provides a value between 0 and 1, where 0 means there is no overlap, and 1 means the prediction result fully matches the ground truth. This metric gives an idea of the extent to which the model's predicted results match the ground truth in the context of an object segmentation task. The higher the IoU value, the better the model performance.

8.2 Dice Coefficient

Dice coefficient is a method to compare the similarity of two different text samples [101]. Dice coefficient is a semimetric version of the Jaccard coefficient. Dice coefficient (also known as F1-score) is used as an evaluation metric to measure the extent to which the segmentation result of the model matches the ground truth. DICE coefficient is defined as twice the number of shared areas (intersection) divided by the total number of areas between the segmentation result and the ground truth. Mathematically, the DICE coefficient formula can be expressed as follows:

$$Dice = \frac{2 \times Intersection}{Total\ Area\ of\ Prediction + Total\ Area\ of\ Ground\ Truth} \quad (2)$$

In the context of U-Net, where the model is used for image segmentation (usually for object segmentation in medical images), "Intersection" is the area where the segmentation results of the model and ground truth match, and "Total Area of Prediction" and "Total Area of Ground Truth" are the total areas of the segmentation results of the model and ground truth, respectively.

8.3 Precision

Precision measures the extent to which the model can correctly identify positives. It is calculated by dividing the number of true positives (correctly identified positives) by the total number of positives identified by the model (true positives plus false positives). In the formula, precision can be explained as follows.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (3)$$

True Positive (TP) is the number of positive instances correctly identified by the model. False Positive (FP) is the number of negative instances incorrectly identified as positive by the model. Precision provides information on how many of the instances identified as positive by the model are actually positive. This metric is important in cases where false positive errors can have significant consequences.

8.4 Recall

Recall, also known as Sensitivity or True Positive Rate, is an evaluation metric that measures the extent to which the model can identify all true positive instances. The recall formula is expressed in the following equation.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (4)$$

Recall gives an idea of the extent to which the model can "remember" or identify true positive instances. This metric is useful in situations where false negative errors can have significant consequences.

8.5 Pixel Accuracy

Pixel accuracy is essentially the number of correctly classified pixels in the resulting segmentation mask [102]. This is probably the simplest metric to evaluate performance, but it may not really address model performance. The problem with the pixel accuracy metric is that it is always biased when there is extreme class imbalance in the data set. It can get ~90% accuracy but the actual qualitative performance will be worse. The pixel accuracy formula is expressed as follows.

$$Pixel\ Accuracy = \frac{Number\ of\ Correctly\ Predicted\ Pixels}{Total\ Number\ of\ Pixels} \quad (5)$$

"Number of Correctly Predicted Pixels" is the number of pixels that have been correctly predicted by the model, and "Total Number of Pixels" is the total number of pixels in the image. Pixel accuracy provides an overview of the extent to which the model was successful in predicting each pixel.

9. CONCLUSION AND FUTURE WORKS

Various previous studies have shown that U-Net as a CNN architecture has great potential in segmenting medical images, especially DFU. The development of variants of U-Net is also an important inspiration in developing automatic segmentation specifically for DFU images. One of the important problems faced by many researchers, especially those dealing with DFU images, is the availability of quality datasets and the need for clinical trials. Datasets also still have a lot of variations in image quality and the challenges of heterogeneity and subjectivity in the annotation process. The creation of a specialised dataset of DFU images is particularly challenging because each ethnicity also has different types and forms of skin lesions.

U-Net as a semantic-based segmentation model must of course be adjusted to the quality of the available dataset. This is made clear by the research results that have different model evaluation values. There are low accuracy evaluation results but there are also evaluation results that have very high accuracy values. The difference in accuracy results is a new idea in developing a new U-Net model that can match a variety of datasets that have different qualities.

Fuzzy logic comes with the hope that it can be applied to the U-Net architecture. Some inspirations for the application of fuzzy logic include its application to the encoder, decoder, skip connection to the final layer of U-Net. The potential of fuzzy logic in U-Net is expected to answer the challenges of various types of DFU image quality variations and also significant differences in accuracy results from several previous studies.

Future research can evaluate the potential use of fuzzy logic in U-Net architecture. This potential is not only developed in standard U-Net, but can be applied to several other types of U-Net variabs. The author is also interested in the application of metaheuristics when fuzzy logic is applied to the U-Net architecture. Metaheuristics are expected to be one of the optimisation techniques of the U-Net fuzzy layer.

REFERENCES

[1] X. Wang *et al.*, "The awareness and determinants of diabetic foot ulcer prevention among diabetic patients: Insights from NHANES (2011–2018)," *Prev Med Rep*, vol. 36, p. 102433, 2023, doi: https://doi.org/10.1016/j.pmedr.2023.102433.

[2] A. V. A. Mariadoss, A. S. Sivakumar, C.-H. Lee, and S. J. Kim, "Diabetes mellitus and diabetic foot ulcer: Etiology, biochemical and molecular based treatment strategies via gene and nanotherapy," *Biomedicine & Pharmacotherapy*, vol. 151, p. 113134, 2022, doi: https://doi.org/10.1016/j.biopha.2022.113134.

[3] G. Balasubramanian, P. Vas, N. Chockalingam, and R. Naemi, "A synoptic overview of neurovascular interactions in the foot," *Front Endocrinol (Lausanne)*, vol. 11, p. 308, 2020.doi: 10.3389/fendo.2020.00308

[4] G. Han and R. Ceilley, "Chronic wound healing: a review of current management and treatments," *Adv Ther*, vol. 34, pp. 599–610, 2017. doi: 10.1007/s12325-017-0478-y

[5] D. Othman, "Negative pressure wound therapy literature review of efficacy, cost effectiveness, and impact on patients' quality of life in chronic wound management and its implementation in the United kingdom," *Plast Surg Int*, vol. 2012, 2012. doi: 10.1155/2012/374398

[6] H. Khalil, M. Cullen, H. Chambers, N. Steers, and J. Walker, "Implementation of a successful electronic wound documentation system in rural Victoria, Australia: a subject of collaboration and community engagement," *Int Wound J*, vol. 11, no. 3, pp. 314–318, 2014. doi: 10.1111/iwj.12041

[7] J. Zhang, C. Mihai, L. Tüshaus, G. Scebba, O. Distler, and W. Karlen, "Wound image quality from a mobile health tool for home-based chronic wound management with real-time quality feedback: Randomized feasibility study," *JMIR Mhealth Uhealth*, vol. 9, no. 7, p. e26149, 2021. doi: 10.2196/26149

[8] P. Sheehan, P. Jones, A. Caselli, J. M. Giurini, and A. Veves, "Percent change in wound area of diabetic foot ulcers over a 4-week period is a robust predictor of complete healing in a 12-week prospective trial," *Diabetes Care*, vol. 26, no. 6, pp. 1879–1882, 2003. doi: 10.2337/diacare.26.6.1879

[9] X. Li *et al.*, "HAL-IA: A Hybrid Active Learning framework using Interactive Annotation for medical image segmentation," *Med Image Anal*, vol. 88, p. 102862, 2023, doi: https://doi.org/10.1016/j.media.2023.102862.

[10] Y. Tang, S. Wang, Y. Qu, Z. Cui, and W. Zhang, "Consistency and adversarial semi-supervised learning for medical image segmentation," *Comput Biol Med*, vol. 161, p. 107018, 2023, doi: https://doi.org/10.1016/j.compbiomed.2023.107018.

[11] H. Zhang *et al.*, "BCU-Net: Bridging ConvNeXt and U-Net for medical image segmentation," *Comput Biol Med*, vol. 159, p. 106960, 2023, doi: https://doi.org/10.1016/j.compbiomed.2023.106960.

[12] J. Cheng *et al.*, "DDU-Net: A dual dense U-structure network for medical image segmentation," *Appl Soft Comput*, vol. 126, p. 109297, 2022, doi: https://doi.org/10.1016/j.asoc.2022.109297.

[13] Z. Wang, J. Zhu, S. Fu, S. Mao, and Y. Ye, "RFPNet: Reorganizing feature pyramid networks for medical image segmentation," *Comput Biol Med*, vol. 163, p. 107108, 2023, doi: https://doi.org/10.1016/j.compbiomed.2023.107108.

[14] W. Baccouch, S. Oueslati, B. Solaiman, and S. Labidi, "A comparative study of CNN and U-Net performance for automatic segmentation of medical images: application to cardiac MRI," *Procedia Comput Sci*, vol. 219, pp. 1089–1096, 2023, doi: https://doi.org/10.1016/j.procs.2023.01.388.

[15] G. Chen, L. Li, J. Zhang, and Y. Dai, "Rethinking the unpretentious U-net for medical ultrasound image segmentation," *Pattern Recognit*, vol. 142, p. 109728, 2023, doi: https://doi.org/10.1016/j.patcog.2023.109728.

[16] R. Gamal, H. Barka, and M. Hadhoud, "GAU U-Net for multiple sclerosis segmentation," *Alexandria Engineering Journal*, vol. 73, pp. 625–634, 2023, doi: https://doi.org/10.1016/j.aej.2023.04.069.

[17] 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 J Biomed Health Inform*, vol. 23, no. 4, pp. 1730–1741, 2019, doi: 10.1109/JBHI.2018.2868656.

[18] N. Rania, H. Douzi, L. Yves, and T. Sylvie, *Semantic segmentation of diabetic foot ulcer images: Dealing with small dataset in dl approaches*, vol. 12119 LNCS. Springer International Publishing, 2020. doi: 10.1007/978-3-030-51935-3_17.

[19] G. Scebba *et al.*, "Detect-and-segment: A deep learning approach to automate wound image segmentation," *Inform Med Unlocked*, vol. 29, no. November 2021, p. 100884, 2022, doi: https://doi.org/10.1016/j.imu.2022.100884.

[20] A. Heras-Tang, D. Valdes-Santiago, A. Leon-Mecias, M. L. B. Diaz-Romañach, and J. A. Mesejo-Chiong, "Diabetic foot ulcer segmentation using logistic regression, DBSCAN clustering and morphological operators," *Electronic Letters on Computer Vision and Image Analysis*, vol. 21, no. 2, pp. 23–39, 2022, doi: https://doi.org/10.5565/REV/ELCVIA.1413.

[21] M. Hu, Y. Zhong, S. Xie, H. Lv, and Z. Lv, "Fuzzy System Based Medical Image Processing for Brain Disease Prediction ," 2021.

[22] M. Cui, K. Li, J. Chen, and W. Yu, "CM-Unet: A Novel Remote Sensing Image Segmentation Method Based on Improved U-Net," *IEEE Access*, vol. 11, pp. 56994–57005, 2023, doi: https://doi.org 10.1109/ACCESS.2023.3282778.

[23] B. Li, T. Yang, and X. Zhao, "NVTrans-UNet: Neighborhood vision transformer based U-Net for multi-modal cardiac MR image segmentation.," *J Appl Clin Med Phys*, vol. 24, no. 3, p. e13908, Mar. 2023, doi: https://doi.org 10.1002/acm2.13908.

[24] R. Sharma, T. Goel, M. Tanveer, S. Dwivedi, and R. Murugan, "FAF-DRVFL: Fuzzy activation function based deep random vector functional links network for early diagnosis of Alzheimer disease," *Appl Soft Comput*, vol. 106, p. 107371, 2021, doi: https://doi.org/10.1016/j.asoc.2021.107371.

[25] M. Kirichev, T. Slavov, and G. Momcheva, "Fuzzy U-Net Neural Network Design for Image Segmentation BT - Contemporary Methods in Bioinformatics and Biomedicine and Their Applications," S. S. Sotirov, T. Pencheva, J. Kacprzyk, K. T. Atanassov, E. Sotirova, and G. Staneva, Eds., Cham: Springer International Publishing, 2022, pp. 177–184.

[26] P. Shah *et al.*, "Wagner's Classification as a Tool for Treating Diabetic Foot Ulcers: Our Observations at a Suburban Teaching Hospital.," *Cureus*, vol. 14, no. 1, p. e21501, Jan. 2022, doi: https://doi.org 10.7759/cureus.21501.

[27] M. Jalilian and S. Shiri, "The reliability of the Wagner Scale for evaluation the diabetic wounds: A literature review," *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 16, no. 1, p. https://doi.org 102369, 2022, doi: https://doi.org/10.1016/j.dsx.2021.102369.

[28] W. D. J. McComb, "Wagner Grading of Diabetic Foot Ulcers and National Pressure Injury Advisory Panel Staging of Pressure Injuries: A Comparison for Clinical Use," *Adv Skin Wound Care*, vol. 36, no. 5, 2023.

[29] A. M. Settembrini and F. Settembrini, "Chapter 28 - Wound care: how to approach venous, arterial and diabetic ulcers," P. Settembrini and A. M. B. T.-V. S. Settembrini, Eds., Academic Press, 2022, pp. 325–335. doi: https://doi.org/10.1016/B978-0-12-822113-6.00010-3.

[30] C. Wang *et al.*, "Fully automatic wound segmentation with deep convolutional neural networks," *Sci Rep*, vol. 10, no. 1, pp. 1–9, 2020, doi: https://doi.org/10.1038/s41598-020-78799-w.

[31] A. Wagh *et al.*, "Semantic Segmentation of Smartphone Wound Images: Comparative Analysis of AHRF and CNN-Based Approaches.," *IEEE Access*, vol. 8, pp. 181590–181604, 2020, doi: https://doi.org/10.1109/access.2020.3014175.

[32] M. Ahsan, S. Naz, R. Ahmad, H. Ehsan, and A. Sikandar, "A Deep Learning Approach for Diabetic Foot Ulcer Classification and Recognition," *Information (Switzerland)*, vol. 14, no. 1, pp. 1–10, 2023, doi: https://doi.org/10.3390/info14010036.

[33] G. Scebba *et al.*, "Detect-and-segment: A deep learning approach to automate wound image segmentation," *Inform Med Unlocked*, vol. 29, p. 100884, 2022, doi: https://doi.org/10.1016/j.imu.2022.100884.

[34] R. Niri, D. Hassan, Y. Lucas, and S. Treuillet, "A Superpixel-Wise Fully Convolutional Neural Network Approach for Diabetic Foot Ulcer Tissue Classification," 2021, pp. 308–320. doi: https://doi.org/10.1007/978-3-030-68763-2_23.

[35] X. Zhao *et al.*, "Fine-Grained Diabetic Wound Depth and Granulation Tissue Amount Assessment Using Bilinear Convolutional Neural Network," *IEEE Access*, vol. 7, pp. 179151–179162, 2019, doi: https://doi.org/10.1109/ACCESS.2019.2959027.

[36] H. N. Huang *et al.*, "Image segmentation using transfer learning and Fast R-CNN for diabetic foot wound treatments," *Front Public Health*, vol. 10, no. 1, 2022, doi: https://doi.org/10.3389/fpubh.2022.969846.

[37] C. Wang *et al.*, "Fully automatic wound segmentation with deep convolutional neural networks," *Sci Rep*, vol. 10, no. 1, p. 21897, 2020, doi: https://doi.org/10.1038/s41598-020-78799-w.

[38] J. P., S. K. B. K., and S. Jayaraman, "Automatic foot ulcer segmentation using conditional generative adversarial network (AFSegGAN): A wound management system," *PLOS Digital Health*, vol. 2, no. 11, p. e0000344, Nov. 2023.

[39] M. Goyal, M. H. Yap, N. D. Reeves, S. Rajbhandari, and J. Spragg, "Fully convolutional networks for diabetic foot ulcer segmentation," in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2017, pp. 618–623. doi: https://doi.org/10.1109/SMC.2017.8122675.

[40] P. N. Thotad, G. R. Bharamagoudar, and B. S. Anami, "Diabetic foot ulcer detection using deep learning approaches," *Sensors International*, vol. 4, p. 100210, 2023, doi: https://doi.org/10.1016/j.sintl.2022.100210.

[41] S. K. Das *et al.*, "Diabetic Foot Ulcer Identification: A Review.," *Diagnostics (Basel)*, vol. 13, no. 12, Jun. 2023, doi: https://doi.org/10.3390/diagnostics13121998.

[42] M. H. Yap *et al.*, "Deep learning in diabetic foot ulcers detection: A comprehensive evaluation," *Comput Biol Med*, vol. 135, p. 104596, 2021, doi: https://doi.org/10.1016/j.compbiomed.2021.104596.

[43] D. Yuan, Y. Liu, Z. Xu, Y. Zhan, J. Chen, and T. Lukasiewicz, "Painless and accurate medical image analysis using deep reinforcement learning with task-oriented homogenized automatic pre-processing," *Comput Biol Med*, vol. 153, p. 106487, 2023, doi: https://doi.org/10.1016/j.compbiomed.2022.106487.

[44] A. Davradou, E. Protopapadakis, M. Kaselimi, A. Doulamis, and N. Doulamis, "Diabetic Foot Ulcers Monitoring by Employing Super Resolution and Noise Reduction Deep Learning Techniques," in *Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments*, in PETRA '22. New York, NY, USA: Association for Computing Machinery, 2022, pp. 83–88. doi: https://doi.org/10.1145/3529190.3529214.

[45] Z. Mao *et al.*, "Deep learning based noise reduction method for automatic 3D segmentation of the anterior of lamina cribrosa in optical coherence tomography volumetric scans," *Biomed Opt Express*, vol. 10, no. 11, pp. 5832–5851, 2019, doi: https://doi.org/10.1364/BOE.10.005832.

[46] N. Uzakkyzy *et al.*, "Image noise reduction by deep learning methods," *International Journal of Electrical and Computer Engineering*, vol. 13, no. 6, pp. 6855–6861, 2023, doi: https://doi.org/10.11591/ijece.v13i6.pp6855-6861.

[47] B. Cassidy *et al.*, "Artificial intelligence for automated detection of diabetic foot ulcers: A real-world proof-of-concept clinical evaluation," *Diabetes Res Clin Pract*, vol. 205, p. 110951, 2023, doi: https://doi.org/10.1016/j.diabres.2023.110951.

[48] J. Huang, G. Liu, and B. Wang, "Semantic Segmentation under a Complex Background for Machine Vision Detection Based on Modified UPerNet with Component Analysis Modules," *Math Probl Eng*, vol. 2020, p. 6903130, 2020, doi: https://doi.org/10.1155/2020/6903130.

[49] L. Alzubaidi *et al.*, "Towards a Better Understanding of Transfer Learning for Medical Imaging: A Case Study," *Applied Sciences*, vol. 10, p. 4523, Jun. 2020, doi: https://doi.org/10.3390/app10134523.

[50] S. Wang *et al.*, "Annotation-efficient deep learning for automatic medical image segmentation," *Nat Commun*, vol. 12, no. 1, p. 5915, 2021, doi: https://doi.org/10.1038/s41467-021-26216-9.

[51] S. Muralidhara, A. Lucieri, A. Dengel, and S. Ahmed, "Holistic multi-class classification & grading of diabetic foot ulcerations from plantar thermal images using deep learning.," *Health Inf Sci Syst*, vol. 10, no. 1, p. 21, Dec. 2022, doi: https://doi.org/10.1007/s13755-022-00194-8.

[52] M. Goyal, N. D. Reeves, S. Rajbhandari, N. Ahmad, C. Wang, and M. H. Yap, "Recognition of ischaemia and infection in diabetic foot ulcers: Dataset and techniques," *Comput Biol Med*, vol. 117, p. 103616, 2020, doi: https://doi.org/10.1016/j.compbiomed.2020.103616.

[53] B. Preim and C. Botha, "Chapter 2 - Acquisition of Medical Image Data," in *Visual Computing for Medicine (Second Edition)*, B. Preim and C. B. T.-V. C. for M. (Second E. Botha, Eds., Boston: Morgan Kaufmann, 2014, pp. 15–67. doi: https://doi.org/10.1016/B978-0-12-415873-3.00002-X.

[54] D. Wade, "Ethics of collecting and using healthcare data.," Jun. 2007, *England*. doi: https://doi.org/10.1136/bmj.39247.679329.80.

[55] S. T. Padmapriya and S. Parthasarathy, "Ethical Data Collection for Medical Image Analysis: a Structured Approach," *Asian Bioeth Rev*, 2023, doi: https://doi.org/10.1007/s41649-023-00250-9.

[56] V. Totten, E. L. Simon, M. Jalili, and H. R. Sawe, "Acquiring data in medical research: A research primer for low- and middle-income countries," *African Journal of Emergency Medicine*, vol. 10, pp. S135–S139, 2020, doi: https://doi.org/10.1016/j.afjem.2020.09.009.

[57] A. Rebinth and M. Kumar S, "Importance of Manual Image Annotation Tools and Free Datasets for Medical Research," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 10, pp. 1880–1885, Jan. 2019.

[58] D. Filko and E. K. Nyarko, "2D/3D Wound Segmentation and Measurement Based on a Robot-Driven Reconstruction System," 2023. doi: https://doi.org/10.3390/s23063298.

[59] Y. Xu and R. Goodacre, "On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning," *J Anal Test*, vol. 2, Oct. 2018, doi: https://doi.org/10.1007/s41664-018-0068-2.

[60] J. Nalepa, M. Marcinkiewicz, and M. Kawulok, "Data Augmentation for Brain-Tumor Segmentation: A Review ," 2019.

[61] N. Rania, H. Douzi, L. Yves, and T. Sylvie, "Semantic Segmentation of Diabetic Foot Ulcer Images: Dealing with Small Dataset in DL Approaches," in *Image and Signal Processing*, A. El Moataz, D. Mammass, A. Mansouri, and F. Nouboud, Eds., Cham: Springer International Publishing, 2020, pp. 162–169.

[62] A. Semma, S. Lazrak, Y. Hannad, M. Boukhani, and Y. El Kettani, "Writer Identification: The Effect of Image Resizing on CNN Performance," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, vol. 46, no. 4/W5-2021, pp. 501–507, Dec. 2021, doi: https://doi.org/10.5194/isprs-Archives-XLVI-4-W5-2021-501-2021.

[63] S. Saponara and A. Elhanashi, "Impact of Image Resizing on Deep Learning Detectors for Training Time and Model Performance BT - Applications in Electronics Pervading Industry, Environment and Society," S. Saponara and A. De Gloria, Eds., Cham: Springer International Publishing, 2022, pp. 10–17.

[64] A. Depeursinge, J. Fageot, and O. S. Al-Kadi, "Chapter 1 - Fundamentals of Texture Processing for Biomedical Image Analysis: A General Definition and Problem Formulation," in *The Elsevier and MICCAI Society Book Series*, A. Depeursinge, O. S. Al-Kadi, and J. R. B. T.-B. T. A. Mitchell, Eds., Academic Press, 2017, pp. 1–27. doi: https://doi.org/10.1016/B978-0-12-812133-7.00001-6.

[65] K. Tran, J. P. Bøtker, A. Aframian, and K. Memarzadeh, "Chapter 6 - Artificial intelligence for medical imaging," A. Bohr and K. B. T.-A. I. in H. Memarzadeh, Eds., Academic Press, 2020, pp. 143–162. doi: https://doi.org/10.1016/B978-0-12-818438-7.00006-X.

[66] J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," *J Big Data*, vol. 6, no. 1, p. 27, 2019, doi: 10.1186/s40537-019-0192-5.

[67] K. Ghosh, C. Bellinger, R. Corizzo, P. Branco, B. Krawczyk, and N. Japkowicz, "The class imbalance problem in deep learning," *Mach Learn*, 2022, doi: https://doi.org/10.1007/s10994-022-06268-8.

[68] S. Rezvani and X. Wang, “A broad review on class imbalance learning techniques,” *Appl Soft Comput*, vol. 143, p. 110415, 2023, doi: https://doi.org/10.1016/j.asoc.2023.110415.

[69] N. Siddique, P. Sidike, C. Elkin, and V. Devabhaktuni, “U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications,” *IEEE Access*, vol. PP, p. 1, Jun. 2021, doi: https://doi.org/10.1109/ACCESS.2021.3086020.

[70] M. Khouy, Y. Jabrane, M. Ameur, and A. Hajjam El Hassani, “Medical Image Segmentation Using Automatic Optimized U-Net Architecture Based on Genetic Algorithm,” 2023. doi: https://doi.org/10.3390/jpm13091298.

[71] M. K. Anbudevi and K. Suganthi, “Review of Semantic Segmentation of Medical Images Using Modified Architectures of UNET,” *Diagnostics*, vol. 12, no. 3064, pp. 1–31, 2022, doi: https://doi.org/10.3390/diagnostics12123064.

[72] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation BT - Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015,” N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds., Cham: Springer International Publishing, 2015, pp. 234–241.

[73] J. Kugelman *et al.*, “A comparison of deep learning U-Net architectures for posterior segment OCT retinal layer segmentation,” *Sci Rep*, vol. 12, no. 1, p. 14888, 2022, doi: https://doi.org/10.1038/s41598-022-18646-2.

[74] A. Khanna, N. D. Londhe, S. Gupta, and A. Semwal, “A deep Residual U-Net convolutional neural network for automated lung segmentation in computed tomography images,” *Biocybern Biomed Eng*, vol. 40, no. 3, pp. 1314–1327, 2020, doi: https://doi.org/10.1016/j.bbe.2020.07.007.

[75] G. Yu, J. Dong, Y. Wang, and X. Zhou, “RUC-Net: A Residual-Unet-Based Convolutional Neural Network for Pixel-Level Pavement Crack Segmentation,” 2023. doi: https://doi.org/10.3390/s23010053.

[76] Q. Xu, Z. Ma, N. HE, and W. Duan, “DCSAU-Net: A deeper and more compact split-attention U-Net for medical image segmentation,” *Comput Biol Med*, vol. 154, p. 106626, 2023, doi: https://doi.org/10.1016/j.compbiomed.2023.106626.

[77] R. Yousef *et al.*, “U-Net-Based Models towards Optimal MR Brain Image Segmentation,” 2023. doi: https://doi.org/10.3390/diagnostics13091624.

[78] M. Z. Alom, C. Yakopcic, M. Hasan, T. M. Taha, and V. K. Asari, “Recurrent residual U-Net for medical image segmentation.,” *J Med Imaging (Bellingham)*, vol. 6, no. 1, p. 14006, Jan. 2019, doi: https://doi.org/10.1117/1.JMI.6.1.014006.

[79] X. Liu, R. Yin, and J. Yin, “Attention V-Net: A Modified V-Net Architecture for Left Atrial Segmentation,” 2022. doi: https://doi.org/10.3390/app12083764.

[80] A. Chaurasia and E. Culurciello, *LinkNet: Exploiting encoder representations for efficient semantic segmentation*. 2017. doi: https://doi.org/10.1109/VCIP.2017.8305148.

[81] J.-X. Cai, T.-J. Mu, Y.-K. Lai, and S.-M. Hu, “LinkNet: 2D-3D linked multi-modal network for online semantic segmentation of RGB-D videos,” *Comput Graph*, vol. 98, pp. 37–47, 2021, doi: https://doi.org/10.1016/j.cag.2021.04.013.

[82] D. Jha, M. A. Riegler, D. Johansen, P. Halvorsen, and H. D. Johansen, “DoubleU-Net: A Deep Convolutional Neural Network for Medical Image Segmentation,” in *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*, 2020, pp. 558–564. doi: https://doi.org/10.1109/CBMS49503.2020.00111.

[83] Z. Liu, J. John, and E. Agu, “Diabetic Foot Ulcer Ischemia and Infection Classification Using EfficientNet Deep Learning Models,” *IEEE Open J Eng Med Biol*, vol. 3, pp. 189–201, 2022, doi: https://doi.org/10.1109/OJEMB.2022.3219725.

[84] D. Kucharski, A. Kostuch, F. Noworolnik, A. Brodzicki, and J. Jaworek-Korjakowska, “DFU-Ens: End-to-End Diabetic Foot Ulcer Segmentation Framework with Vision Transformer Based Detection BT - Diabetic Foot Ulcers Grand Challenge,” M. H. Yap, C. Kendrick, and B. Cassidy, Eds., Cham: Springer International Publishing, 2023, pp. 101–112.

[85] S. M. Shah, A. Rizwan, G. Atteia, and M. Alabdulhafith, “CADFU for Dermatologists: A Novel Chronic Wounds & Ulcers Diagnosis System with DHuNeT (Dual-Phase Hyperactive UNet) and YOLOv8 Algorithm,” 2023. doi: https://doi.org/10.3390/healthcare11212840.

[86] K. Sanjar, O. Bekhzod, J. Kim, J. Kim, A. Paul, and J. Kim, “Improved U-Net: Fully Convolutional Network Model for Skin-Lesion Segmentation,” 2020. doi: https://doi.org/10.3390/app10103658.

[87] B. Chudasama, N. Ovaskainen, J. Tamminen, N. Nordbäck, J. Engström, and I. Aaltonen, "Automated mapping of bedrock-fracture traces from UAV-acquired images using U-Net convolutional neural networks," *Comput Geosci*, vol. 182, p. 105463, 2024, doi: https://doi.org/10.1016/j.cageo.2023.105463.

[88] J. Arnal and L. Súcar, "Fast Method Based on Fuzzy Logic for Gaussian-Impulsive Noise Reduction in CT Medical Images," 2022. doi: https://doi.org/10.3390/math10193652.

[89] Y.-P. Wang *et al.*, "Use of U-Net Convolutional Neural Networks for Automated Segmentation of Fecal Material for Objective Evaluation of Bowel Preparation Quality in Colonoscopy," *Diagnostics*, vol. 12, p. 613, Mar. 2022, doi: https://doi.org/10.3390/diagnostics12030613.

[90] H. Lu, Y. She, J. Tie, and S. Xu, "Half-UNet: A Simplified U-Net Architecture for Medical Image Segmentation ," 2022.

[91] T. Xiang *et al.*, "Towards bi-directional skip connections in encoder-decoder architectures and beyond," *Med Image Anal*, vol. 78, p. 102420, 2022, doi: https://doi.org/10.1016/j.media.2022.102420.

[92] M. M. Hasan, M. M. Hossain, M. M. Rahman, A. K. M. Azad, S. A. Alyami, and M. A. Moni, "FP-CNN: Fuzzy pooling-based convolutional neural network for lung ultrasound image classification with explainable AI," *Comput Biol Med*, vol. 165, p. 107407, 2023, doi: https://doi.org/10.1016/j.compbiomed.2023.107407.

[93] D. Kucharski, A. Kostuch, F. Noworolnik, A. Brodzicki, and J. Jaworek-Korjakowska, "DFU-Ens: End-to-End Diabetic Foot Ulcer Segmentation Framework With Vision Transformer Based Detection," in *Diabetic Foot Ulcers Grand Challenge: Third Challenge, DFUC 2022, Held in Conjunction with MICCAI 2022, Singapore, September 22, 2022, Proceedings*, Berlin, Heidelberg: Springer-Verlag, 2023, pp. 101–112. doi: https://doi.org/10.1007/978-3-031-26354-5_9.

[94] X. Chen, D. Li, P. Wang, and X. Yang, "A Deep Convolutional Neural Network with Fuzzy Rough Sets for FER," *IEEE Access*, vol. 8, pp. 2772–2779, 2020, doi: https://doi.org/10.1109/ACCESS.2019.2960769.

[95] C. Guan, S. Wang, and A. W. C. Liew, "Lip Image Segmentation Based on a Fuzzy Convolutional Neural Network," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 7, pp. 1242–1251, 2020, doi: https://doi.org/10.1109/TFUZZ.2019.2957708.

[96] T.-D.-T. Phan, S.-H. Kim, H.-J. Yang, and G.-S. Lee, "Skin Lesion Segmentation by U-Net with Adaptive Skip Connection and Structural Awareness," 2021. doi: https://doi.org/10.3390/app11104528.

[97] T. Sharma, N. K. Verma, and S. Masood, "Mixed fuzzy pooling in convolutional neural networks for image classification," *Multimed Tools Appl*, vol. 82, no. 6, pp. 8405–8421, 2023, doi: https://doi.org/10.1007/s11042-022-13553-0.

[98] C. J. Lin and J. Y. Jhang, "Intelligent Traffic-Monitoring System Based on YOLO and Convolutional Fuzzy Neural Networks," *IEEE Access*, vol. 10, pp. 14120–14133, 2022, doi: https://doi.org/10.1109/ACCESS.2022.3147866.

[99] J. Yu *et al.*, "Learning Generalized Intersection Over Union for Dense Pixelwise Prediction," in *Proceedings of the 38th International Conference on Machine Learning*, M. Meila and T. Zhang, Eds., in Proceedings of Machine Learning Research, vol. 139. PMLR, 2021, pp. 12198–12207.

[100] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, *Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression*. 2019.

[101] R. R. Shamir, Y. Duchin, J. Kim, G. Sapiro, and N. Harel, "Continuous Dice Coefficient: a Method for Evaluating Probabilistic Segmentations," *bioRxiv*, p. 306977, Jan. 2018, doi: https://doi.org/10.1101/306977.

[102] F. Shafait, D. Keysers, and T. Breuel, *Pixel-Accurate Representation and Evaluation of Page Segmentation in Document Images*, vol. 1. 2006. doi: https://doi.org/10.1109/ICPR.2006.934.