Classification of diabetic foot ulcer using convolutional neural network (CNN) in diabetic patients

Mawaddah Harahap¹ ¹, Sai Kumarani Anjelli², Widy Anggun M. Sinaga³, Ryan Alward⁴, Junio Fegri Wira Manawan⁵, Amir Mahmud Husein⁶

¹-⁶Universitas Prima Indonesia
¹-⁶Jalan Sampul, Sei Putih Barat, Medan 20118, Indonesia
*Corresponding email: amirmahmud@unprimdn.ac.id

Received 29 June 2022, Revised 26 July 2022, Accepted 27 July 2022

Abstract — The image of chronic wounds on human skin tissue has the similar look in shape, color, and size to each other even though they are caused by different diseases. A diabetic ulcer is a condition where peripheral arterial blood vessels are disrupted due to hyperglycemia in people with diabetes mellitus. This research aimed to analyze the accuracy of the Convolutional Neural Network algorithm in classifying diabetic ulcer disease with a transfer learning approach based on the appearance of the image of the wound on the sole in people with diabetes mellitus. By applying the transfer learning approach, the results showed that the ResNet152V2 model achieved the best accuracy value of 0.993 (99%), the precision of 1.00, recall of 0.986, F1-Score of 0.993, and Support of 72. Therefore, the ResNet152V2 model was highly considered for classifying diabetic ulcers in patients with diabetes mellitus.

Keywords – classification, convolutional neural network, diabetic foot ulcer, diabetes mellitus, transfer learning

I. INTRODUCTION

Humans have the similar look in shape, color, and size of wounds to each other even though they are caused by different diseases [1]. Especially if the images taken are in hard-to-reach parts of the body such as the toes and soles. A metabolic disorder characterized by symptoms of hyperglycemia causing both insulin secretion and performance to be abnormal [2]. The hyperglycemic condition of people with diabetes mellitus can increase the risk of complications for several other diseases such as retinopathy (damage to blood vessels in the eye that can lead to blindness), cardiovascular, nephropathy, and peripheral neuropathy as a cause of diabetic ulcers [3].

A diabetic ulcer is a condition where the peripheral arterial blood vessels are disrupted due to hyperglycemia in DM patients [3]. Disturbances in the blood vessels will cause wounds accompanied by infections of the skin tissue of the feet so it requires treatment time of months or even years [4]. In addition, the appearance of diabetic ulcers has an almost similar look to wounds from different diseases. Indeed, it makes it more difficult for medical personnel to diagnose patients. If the wound in diabetic ulcer patients is not treated properly, the patient can experience amputation which causes physical disability, decreased quality of life, and even death [5]. In recent years, there have been many researchers who have studied deep learning methods that are considered effective for recognizing diabetic ulcers [6]–[21].

Convolutional Neural Network (CNN) is a deep learning algorithm that is widely used to process image data because it has a high network depth [12]. Before classifying objects in the image of the wound on the sole, it is necessary to do image processing. One of the deep learning algorithms that can be used in image processing is the CNN algorithm [22]. There are several stages in the CNN, starting from the convolution stage by using a certain size kernel into an image. Then proceed with the activation function of the Rectifier Liner Unit which will later be fully connected to the neural network. The result of this fully connected network will be the output class [23].

In research by Cassidy, et al. [24], an evaluation was
carried out on the application of the CNN algorithm with various frameworks for classifying diabetic ulcers in various related studies. Cruz-Vega et al. [25], in their research proposed a new model called DFTNet to classify normal (sheat) and abnormal (diabetic ulcers) wounds with a transfer learning approach while in the research by Galdran, et al. [26], CNN was trained with the SAM optimizer to obtain better accuracy in differentiating between ulcers and other wounds.

The implementation of the CNN algorithm to classify diabetic ulcers has been proven to be accurate [27]–[39]. To provide significant results, the researcher proposes a transfer learning approach to be applied to the CNN algorithm for the classification of diabetic ulcers and non-diabetic ulcers.

II. RESEARCH METHOD

This research was experimental research. In this experiment, the researcher wanted to analyze the accuracy of the CNN algorithm in classifying the image of the soles in patients with diabetic ulcers and healthy feet as shown in Fig. 1.

A. The Proposed Model

This research adapted a network of transfer learning models such as VGG19, MobileNetV2, InceptionResNetV2, ResNet50V2, ResNet101V2, and ResNet152V2, to classify images of normal (healthy) and abnormal feet (diabetic ulcers).

B. Dataset

The image test data was obtained from the online dataset listed on the dataset provider platform https://www.kaggle.com/laithjj/diabetic-foot-ulcer-dfu. Where the dataset consisted of the original image dataset of healthy feet and feet affected by diabetic ulcers, extracted from the original image dataset, images used for training models, and for transfer learning purposes. The image dataset contained 543 images of healthy feet and 512 images of diabetic ulcers.

C. Work Procedure

In this research, the work procedure for processing images with the CNN algorithm can be done through the following steps:

1) Image test data from the DFU dataset were grouped into each folder with normal and abnormal labels. The two folders were then combined and divided into 547 training image sets, 171 test sets, and 137 validation sets. The classification results would be labeled with the number 1 meaning Abnormal and 0 representing the Normal label.

2) Pre-processing of image test data was carried out by rotating, adding contrast, and removing noise in the image.

3) The next step was to input images for processing in order to classify patients with diabetic ulcers using transfer learning models, namely VGG19, MobileNetV2, InceptionResNetV2, and Resnet (ResNet101V2, ResNet152V2, ResNet50V2).

4) Connected Layer as a classifier with feature extraction and evaluation of the probability of objects in the image test data to obtain the final result of model performance.

III. RESULT

The results of the research included the process of data preparation, testing, and discussion of the proposed model for the classification of diabetic ulcers.

A. Data Preparing

The data prepared for testing was sourced from the online dataset dfu-dataset which can be accessed via https://www.kaggle.com/laithjj/diabetic-foot-ulcer-dfu. This dataset consisted of hundreds of pictures of the patients’ feet with diabetic ulcers and healthy feet to be combined in the NORMAL and ABNORMAL folders.
Classification of diabetic foot ulcer using convolutional neural network (CNN) in diabetic patients

Evaluation of model performance was conducted by calculating accuracy, precision, recall, and $F_1$ score ($F_1$-Score).

\[
\text{Accuracy} = \frac{(TN + TP)}{(TN + TP + FN + FP)} \tag{1}
\]

\[
\text{Precision} = \frac{TP}{(TP + FP)} \tag{2}
\]

\[
\text{Recall} = \frac{TP}{(TP + FN)} \tag{3}
\]

\[
F_1 - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}
\]

Where $TP$ was True Positive, $TN$ = True Negative, $FP$ = False Positive, and $FN$ = False Negative. The accuracy of diabetic ulcer classification results against the proposed model was obtained from the value of accuracy, precision, recall, and $F_1$-Score after testing.

B. Testing

The platform to accommodate the testing in this research was Google Collab as a means of making Python-based applications and storing datasets for research purposes using the Google Drive platform. The existence of architectural differences in transfer learning techniques required performance analysis in order to classify people with diabetic ulcers and normal patients.

The pre-processing stage was the first stage carried out on the image by changing the image size (scaling) to 224 x 224 pixels, rotating the image, and removing colored spots (noise) to clarify the focus on the image test data. All models used in this research were trained with the Adams optimizer. The results of the evaluation of the overall model proposed in this research can be seen in Fig. 3 to Fig. 8.

Fig. 3 to Fig. 8 show the model that has been trained during the test to classify diabetic ulcers based on the image of the wound on the sole with two labels, namely, label 0 for Normal and label 1 for Abnormal.

Observations were conducted on loss accuracy, training, testing, and validation with a batch size of 15, and setting the epoch value in the number 100 to 500. In the InceptionNetV2, MobileNetV2, ResNet50V2, ResNet101V2, and ResNet152V2 models, there was a change in the accuracy value that increased from each epoch. VGG19 produced fluctuating accuracy for each variant of the epoch value.

Based on the results of the confusion matrix of all

![Fig. 2. DPUS dataset.](image-url)

![Fig. 3. Training loss and accuracy result using VGG19 model.](image-url)

![Fig. 4. Training loss and accuracy result using InceptionResNetV2 model.](image-url)

![Fig. 5. Training loss and accuracy result using MobileNetV2 model.](image-url)
Classification of diabetic foot ulcer using convolutional neural network (CNN) in diabetic patients

Table 1. Accuracy Result’s Comparison

<table>
<thead>
<tr>
<th>Model</th>
<th>Label</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-Score</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>VGG19</td>
<td>0</td>
<td>0.96</td>
<td>0.94</td>
<td>0.98</td>
<td>0.96</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.98</td>
<td>1.00</td>
<td>0.98</td>
<td>0.97</td>
<td>81</td>
</tr>
<tr>
<td>Resnet50V2</td>
<td>0</td>
<td>0.97</td>
<td>0.97</td>
<td>1.00</td>
<td>0.97</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.97</td>
<td>1.00</td>
<td>0.98</td>
<td>0.97</td>
<td>87</td>
</tr>
<tr>
<td>Resnet101V2</td>
<td>0</td>
<td>0.98</td>
<td>1.00</td>
<td>1.00</td>
<td>0.98</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.97</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>86</td>
</tr>
<tr>
<td>Resnet152V2</td>
<td>0</td>
<td>0.99</td>
<td>1.00</td>
<td>0.99</td>
<td>0.99</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>79</td>
</tr>
<tr>
<td>InceptionResNetV2</td>
<td>0</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>71</td>
</tr>
<tr>
<td>MobileNetV2</td>
<td>0</td>
<td>0.97</td>
<td>0.96</td>
<td>0.96</td>
<td>0.97</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.97</td>
<td>87</td>
</tr>
</tbody>
</table>
models that are classified to the two labels, it can be seen in Fig. 9 to Fig. 14 that each model produced a good level of accuracy in classifying Normal legs and Abnormal legs (diabetic ulcers) despite the difference in results. The VGG19 model can classify 89 abnormal images (diabetic ulcers) with one detection error while from the normal image 76 successfully classified and five detection errors. Then the InceptionResNetV2 model can recognize 77 abnormal images and 74 normal images without detection errors. In the MobileNetV2 model, there is an increase in image recognition on abnormal labels by 83 images with one detection error and one normal labels by 84 images with three detection errors.

The ResNet50V2 model is able to classify abnormal images better than the MobileNetV2 model, which is 84 images without detection errors. As for normal images, there are 86 images that were successfully classified with two detection errors. In contrast to ResNet50V2, ResNet101V2 has no detection error when recognizing 86 normal images but in abnormal images, there are 83 images that were successfully recognized with two detection errors. Unlike the previous two ResNet models, ResNet152V2 did not produce a detection error when recognizing 95 images of normal feet and 76 abnormal feet.

Based on the test parameters used to measure the performance of each model in classifying diabetic ulcers, it showed that the Resnet152V2 model was superior to other models with an percentage of accuracy was 0.993 (99%), precision was 1.00, recall was 0.986, \(F_1\)-Score was 0.993 and Support value was 72.

IV. DISCUSSION

In research [1] classifying ulcers by adapting the transfer learning method, namely VGG19 on the CNN algorithm. Tests were carried out on 300 images and 252 images from different dataset sources which were tested only with 100 epochs to obtain different accuracy results, with an average accuracy value of 85%, precision of 32%, and recall of 75%. In this research, better accuracy was obtained for a larger dataset, namely a dataset with a total of 300 images.

The same thing was also applied in research [40], but this research aimed to detect covid and non-covid diseases with the CNN approach through feature extraction of the transfer learning model. Several models applied in this research consisted of VGG19, InceptionNetV2, MobileNetV2, ResNet50V2, ResNet101V2, ResNet152V2. This model was applied to the covid dataset which consisted of 352 training data, 100 testing data, and 88 validation data. From the results of experiments conducted on Google Colab, the ResNet50V2 model proved to have a better accuracy rate than other models with an accuracy rate of 0.95 (95%) Precision 0.96, Recall 0.973, \(F_1\)-Score 0.966, and Support of 74.

The model in the research [40] was adopted in this research and applied to a diabetic ulcer dataset containing 543 images of healthy feet and 512 images of feet affected by diabetic ulcers. Where later this image test data was divided into 547 training sets, 171 test sets, and 137 validation sets. After testing with variations of epoch values from 100 to 500, it was found that the Resnet152V2 model was superior with an accuracy value of 0.993 (99%), precision of 1.00, recall of 0.986, \(F_1\)-Score of 0.993, and Support of 72.
The differences in the model result used were due to differences in epoch values, and a large amount of test data that affected the resulting accuracy.

V. CONCLUSION
The transfer learning model consists of several models, namely the VGG19, MobileNetV2, InceptionResNetV2, ResNet50V2, ResNet101V2, and ResNet152V2 models proposed in this research for the classification of diabetic ulcers based on foot images. The stages of testing were carried out on a collection of datasets grouped into training, testing, and validation data. Furthermore, the pre-processing stage was carried out for testing purposes on the Google Collab platform. The experimental results showed that the ResNet152V2 model achieved the best accuracy value of 0.993 (99%) with a precision of 1.00, recall of 0.986, F1-Score of 0.993, and support of 72. While VGG19 produced the lowest accuracy value due to differences in epoch values and the large amount of test data that affected the value of resulting accuracy. Therefore, the ResNet152V2 model was highly considered to classify diabetic ulcer disease in patients with diabetes mellitus. As a future study, we need to try different transfer learning models to classify the wound images into more classes with a larger dataset.

REFERENCES
Classification of diabetic foot ulcer using convolutional neural network (CNN) in diabetic patients


