An Efficient Ensemble of Deep Neural Networks for Detection and Classification of Diabetic Foot Ulcers Images

Basabi Chakraborty
Research and Regional Cooperation Division
Iwate Prefectural University, Takizawa, Japan
Computer Science and Technology
Madanapalle Institute of Technology and Science
AP, India
email: basabi@iwate-pu.ac.jp

Suma Sailaja Nakka
Software and Information Science
Iwate Prefectural University
Takizawa, Japan
email: g231t026@s.iwate-pu.ac.jp

Takahisa Sanada
Software and Information Science
Iwate Prefectural University
Takizawa, Japan
email: takahisa_s@iwate-pu.ac.jp

***Abstract*—Classification of Diabetic Foot Ulcers (DFU) wounds using computerized methods is becoming an important research area due to development of machine learning and deep learning algorithms for image classification. In this work an efficient ensemble of several deep neural networks has been proposed for classification of DFU images. Simulation experiments with publicly available Diabetic Foot Ulcers Grand Challenge (DFUC 2021) data set has been done to justify the proposal. The performance of the ensemble has been studied and it is found that the ensemble produced a classification accuracy of 91% with a reasonable computational cost which is considered higher compared to the existing approaches.**

***Index Terms*—Diabetic foot ulcers classification, deep neural network, ensemble classifier**

I. Introduction

Deep Neural Networks (DNN) are increasingly used in medical image analysis as a part of the development of computer aided diagnosis systems. Automatic detection and classification of Diabetic Foot Ulcers (DFU) wounds using deep learning tools is one of such applications. DFU is one of the major complications of diabetes. DFU with infection and ischemia can be a serious threat to the patient, leading to death. Early detection and classification of wounds according to the presence of infection, ischemia or both is needed for successful treatment. A comprehensive assessment of several techniques for detection of DFU based on Diabetic Foot Ulcers Grand Challenge (DFUC) 2020 dataset is reported by Yap et al. [1]. The DFUC 2020 classification results reported in various literature so far are summarized by Zhang et al. [2]. Most of the research works on binary classification of DFUC 2020 by deep networks are based on Convolutional Neural Network (CNN) architecture and its several variants. Goyal et al. proposed DFUNet based on CNN [3] and an ensemble of CNN and SVM [4] for binary classification of normal and DFU images and infection vs non-infection, ischemia vs non-ischemia, respectively. Xu et al. [5] used vision transformer model and Das et al. [6] used ResNet for binary classification of infection vs non-infection and ischemia vs non-ischemia classes. Diabetic Foot Ulcers Grand Challenge (DFUC) 2021 results are summarized by Cassidy et al. [7]. DFUC 2021 data set contains four classes of DFU wounds i.e., with infection, with ischemia, with both and none. Several deep learning models and their ensembles are proposed by challengers and the best F1 score on test data were reported as 0.63%.

In this work, several deep neural network models are used for four class classification task of DFUC 2021 data set. The performance of each model has been assessed using several metrics like classification accuracy, precision, recall and F1 score. Finally, an ensemble of top 3 individual neural network models is proposed for the classification task and the performance study by simulation experiments has been done. In section 2, a brief description of the data set and deep neural networks are presented. Section 3 describes the simulation experiments and results followed by section 4 containing conclusion and future work.

II. Data set and Deep Neural Networks

A. Data set

The Diabetic Foot Ulcers data set (DFUC2021) reported by Yap et al. [8] is a pathology analysis data set focusing on infection and ischemia. The final release of DFUC2021 consists of 15,683 DFU patches, with 5,955 training, 5,734 for testing, and 3,994 unlabeled DFU patches. The data set consists of images of four distinct classes which are infection, ischemia, both infection and ischemia, none. Among 5955 training samples, 2,552 samples are without ischemia and infection, 2,555 samples are infection only, and 621 samples are both infection and ischemia, 227 samples are ischemia only. As the number of samples in 4 classes are not balanced, image augmentation techniques (oversampling) are used to increase the number of samples in the classes having less number of samples. Simulation experiment is performed on the both imbalanced and balanced data sets of DFU images.

B. Deep Neural Networks

In this study, popular deep neural networks for various image classification tasks are used. They are VGG16, VGG19,

 ISBN: 978-1-68558-056-8

TABLE I
RESULTS FOR IMBALANCED DATA

Model	Accuracy	Precision	F1 Score	Time
VGG16	0.69	0.50	0.52	20:12
VGG19	0.70	0.54	0.57	15:50
ResNet50	0.53	0.35	0.34	23:22
EfficientNetB0	0.42	0.18	0.25	29:08

ResNet50, DenseNet121, InceptionV3, EfficientNetB0 and EfficientNetB7.

III. SIMULATION EXPERIMENTS AND RESULTS

A. Simulation Experiments

Simulation experiments are done with training set images. All the images are used as RGB images and were resized to a size of $224 \times 224 \times 3$ and normalized. All the network models used in this study are initially pretrained with ImageNet data set. The pre-processed DFU images are then fed to different network models with ImageNet weights. Each model is trained for 20 epochs and batch size of 32 for both imbalanced and balanced data using train-test-validation-split of 70%,20%,10% for training, testing and validation of images respectively. Oversampling of minority classes of the data set is done to increase the number of samples in minority classes equal to majority sample classes. Several image processing techniques like rotating the training images by random rotation angles, horizontal and vertical flips, contrast and brightness, shearing, and zooming in and out of the images have been done as oversampling. Finally, 5-fold Stratified K-fold Cross Validation with k=5 has been done with Stochastic Gradient Descent (SGD) as optimizer for 50 epochs for the balanced data set. To avoid the overfitting of the models we used dropout rate of 50% in all the layers.

B. Simulation Results

Table I represents the performance of some of the models in terms of classification accuracy, precision, F1 score and computational time in min:sec with original data set without balancing with a train-test-validation-split of 70%,20%,10%. VGG19 seems to be the best model according to classification accuracy. After using balanced data sets, the classification accuracy increased to 0.76 for VGG16, 0.79 with VGG19, 0.61 for ResNet50, 0.59 with ResNet101 and 0.54 for EfficientNetB0.

Table II represents the performance of all the models in terms of classification accuracy, Precision, F1score and computational time in hours:min after fine tuning of the models and using Stratified 5-fold Cross Validation technique with 50 epochs over balanced data set. It is found that the top three performing deep network models are DenseNet121, InceptionV3 and VGG19 respectively. An ensemble of the above three models is proposed in this study in order to achieve better results. The initial layers of individuals models are locked with weights pretrained by ImageNet, the last layers are trained by training set followed by the full connected layers for averaging the output prediction of the three independent models. Finally, a softmax activation layer is used for the output class. The classification performance of the ensemble model is reported in the last line of Table II. It is found that the classification accuracy of the ensemble model is higher than the individual models.

TABLE II
RESULTS FOR BALANCED DATA

Model	Accuracy	Precision	F1 Score	Time
VGG16	0.779	0.794	0.776	3:54
VGG19	0.842	0.844	0.841	3:20
ResNet50	0.785	0.786	0.785	3:55
DenseNet121	0.891	0.891	0.891	3:06
EfficientNetB7	0.544	0.469	0.503	3:57
EfficientNetB0	0.433	0.242	0.276	3:58
InceptionV3	0.863	0.865	0.863	3:17
Ensemble	**0.916**	0.917	0.916	4:04

IV. CONCLUSION AND FUTURE WORK

Automatic classification of DFU wounds help doctors in early detection of severity of the disease and save time for treatment. This paper examines several deep neural network models for their effectiveness in four class classification of DFUC 2021 data set and an ensemble of top ranking models is proposed. It has been found that the classification accuracy of the ensemble model is higher than the individual models and reported research works in this problem. But the computational cost is higher than the individual models. The model should be tested for other DFU data sets and compared to other existing high performing models as the future work.

ACKNOWLEDGMENT

This research was partly funded by Japan Society of Promotion of Science (JSPS) KAKENHI Grant Number JP 20K11939.

REFERENCES

[1] M. H. Yap et al., “Deep learning in diabetic foot ulcers detection: A comprehensive evaluation,” Computers in Biology and Medicine, Vol. 135, 104596, Aug. 2021.

[2] J. Zhang et al., ”A comprehensive review of methods based on deep learning for diabetes-related foot ulcers”, Front Endocrinol (Lausanne). 2022 Aug 8;13:945020.

[3] M. Goyal, et.al.,”Dfunet: Convolutional neural networks for diabetic foot ulcer classification. IEEE Trans Emerg Topic Comput Intell vol.4(5), pp. 728– 739, 2018.

[4] M. Goyal, et.al. “Recognition of ischaemia and infection in diabetic foot ulcers: Dataset and techniques,” Computers in Biology and Medicine, vol. 117, no. 103616, 2020.

[5] Y. Xu, et.al., ”Classification of diabetic foot ulcers using class knowledge banks”, Front Bioengineer Biotechnol, Vol 9, 2021. doi:10.3389/fbioe.2021.811028

[6] S. K. Das, P. Roy and A. K. Mishra, ”Recognition of ischaemia and infection indiabetic foot ulcer: a deep convolutional neural network based approach. Int J. Imaging Syst Technol Vol. 32(1),pp.192–208, 2022.

[7] B. Cassidy et.al., ”Diabetic Foot Ulcer Grand Challenge 2021: Evaluation and Summary”, Diabetic Foot Ulcers Grand Challenge: Second Challenge, DFUC 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings. Springer-Verlag, Berlin, Heidelberg.

[8] M. H. Yap, B. Cassidy, J. M. Pappachan, C. O’Shea, D. Gillespie and N. D. Reeves, “Analysis Towards Classification of Infection and Ischaemia of Diabetic Foot Ulcers,” arXiv:2014.03068, 2021.