IEICE TRANSACTIONS

on Information and Systems

This advance publication article will be replaced by
the finalized version after proofreading.

EfficientNet Empowered by Dendritic Learning for Diabetic Retinopathy

Zeyuan JU[†], **Zhipeng LIU**[†], **Yu GAO**[†], **Haotian LI**[†], **Qianhang DU**[†], **Kota YOSHIKAWA**[†], *Nonmembers,* *and* **Shangce GAO**[†a)], *Member*

SUMMARY Medical imaging plays an indispensable role in precise patient diagnosis. The integration of deep learning into medical diagnostics is becoming increasingly common. However, existing deep learning models face performance and efficiency challenges, especially in resource-constrained scenarios. To overcome these challenges, we introduce a novel dendritic neural efficientnet model called DEN, inspired by the function of brain neurons, which efficiently extracts image features and enhances image classification performance. Assessments on a diabetic retinopathy fundus image dataset reveal DEN's superior performance compared to EfficientNet and other classical neural network models.
key words: deep learning, convolutional neural network, diabetic retinopathy, dendritic neuron model

1. Introduction

Diabetes is a metabolic disease characterized by elevated levels of sugar in the blood and urine. Its onset mechanism usually involves insufficient insulin secretion or insulin resistance, leading to ineffective utilization of glucose, resulting in high blood sugar levels. Meanwhile, diabetes not only affects blood sugar levels but also gives rise to a range of severe complications, one of which is diabetic retinopathy (DR) [1]. This condition directly impacts a patient's vision and eye health. Although it may be asymptomatic in its early stages, it can lead to blindness without timely monitoring and intervention. Therefore, eye examinations become crucial for preventing complications.

In the medical domain, the conventional approaches to diagnosing DR necessitate manual examinations. This procedure entails a meticulous evaluation of retinal images by a seasoned and proficient specialist. These examinations exhibit constraints and susceptibility to specific factors such as noise, lighting disparities, and variations in photography equipment. Therefore, effectively extracting features from datasets under these circumstances can be rather demanding [2].

Deep learning methods have been widely used in medical diagnostic problems with great success. They are less affected by environmental factors and can alleviate the difficulty of extracting effective features from large amounts of data. However, existing deep learning models still face some performance and efficiency challenges.

MobileNet [3] has lower accuracy and may perform poor than ResNet50 [4] in tasks that require high accuracy. ResNet50 and DenseNet161 usually achieve high accuracy in image classification, object detection, and semantic segmentation tasks, but they may not be suitable for small datasets [5]. Consequently, researchers have been continuously exploring methods to enhance model performance and efficiency, with a focus on aspects like model size, structure, and characteristics.

Many traditional models have sought to simplify the intricate biological mechanisms at play, providing highly distilled, phenomenological representations of the input-output characteristics of neurons [6]. The McCulloch-Pitts neuron, initially proposed by McCulloch and Pitts, represents a streamlined abstraction of a neuron [7]. A pivotal concept in contemporary deep learning technology is the "perceptron" [8]. However, it is crucial to acknowledge that the fundamental operation of the perceptron entails the linear summation of inputs and the establishment of output thresholds. This simplistic approach overlooks the temporal intricacies inherent in non-linear synaptic integration processes and the true nature of neuronal output.

Gao et al. proposed a dendritic neural model (DNM) which inspired in biological neurons to solve nonlinear problems [9]. This in-depth research has led to its increased utilization for enhancing the performance of medical image classification and segmentation networks [10] [11]. Moreover, its application has expanded from the real-valued domain to the complex-valued domain, with superior performance in various tasks compared to other models [12]. We introduced a novel dendritic efficient neural model called DEN, which incorporates dendritic neural model (DNM) into EfficientNet [13]. This model draws inspiration from the workings of brain neurons, combining efficient image feature extraction with non-linear information processing to enhance performance in image classification tasks. The experimental results suggest that the utilization of the DEN model for DR classification in medical eye images has yielded promising results. Compared to five classical models, the DEN model has performs a notable performance enhancement, with an increase in accuracy by 1.8%. This outcome underscores the potential and superiority of the DEN model in addressing challenges related to medical image classification, providing novel solutions for supporting diagnostics in the medical domain. The main contributions of this paper are as follows:

1) We introduce a novel DEN model, which amalgamates

†Faculty of Engineering, University of Toyama, Toyama 930-8555, Japan.
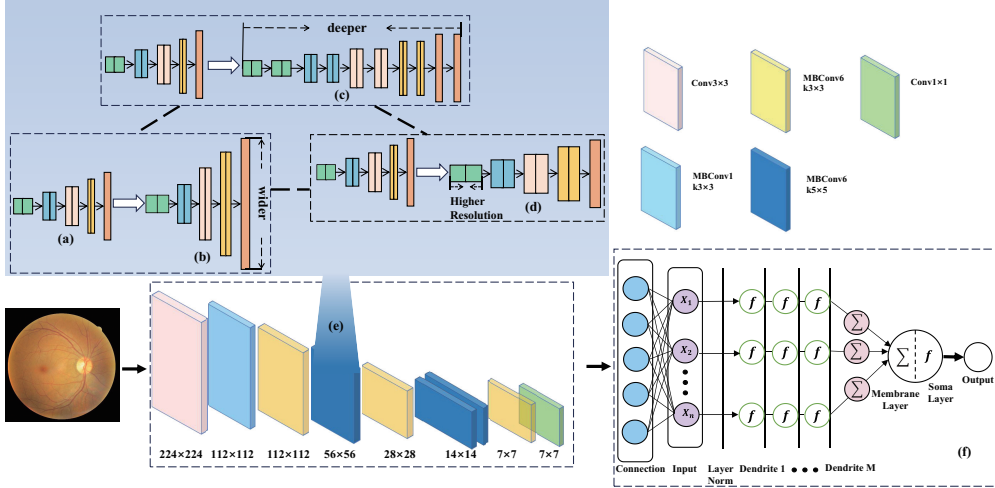a) E-mail: gaosc@eng.u-toyama.ac.jp

Fig. 1: **The DEN structured diagram.** (a) represents a baseline network example. (b)-(d) are indicative of traditional scaling techniques, individually boosting a specific aspect of network width, depth, and resolution. (e) exemplifies the compound scaling approach, scaling all three dimensions with a fixed ratio uniformly. (f) illustrates the structured diagram of DNM.

the non-linear attributes of dendritic neurons with the robust feature extraction capabilities of deep learning, offering a fresh perspective for medical image classification and various application domains.

2) Experimental results validate the superior performance of the proposed DEN model compared to other baseline models, particularly in the classification of medical diabetic retinopathy.

2. Dendritic neural efficient model

In this study, we propose DEN, aimed at enhancing the performance and efficiency in medical image classification tasks. The network structure of DEN consists of two main components: the image feature extraction component and feature classification component. The former is capable of efficiently extracts valuable features from input images. The latter is designed to handle complex non-linear relationships, enabling more flexible feature transformations and weightings, allowing it to adapt to various data types. The overall framework of DEN is depicted in Fig. 1.

Initially, the input data is standardized to enhance the training efficiency. Next, the data is fed into the Efficient-Net network for feature extraction of diabetic eye corneal disease. Furthermore, a feature mapping module, in order to enhance the mapping of feature vectors to prediction outcomes within the DNM, we opted for a more suitable input channel number, set at 128. This feature map encompasses edge details, textures, as well as higher-level semantic information from the images. Subsequently, these insightful features are relayed to the dendritic network, which comprises four biomimetic layers (synapse layer, dendritic layer, membrane layer, and soma layer). These layers engage in in-depth analysis and classification of the extracted features, effectively simulating the information processing

mechanisms found within neurons at each level.

2.1 Image Feature Extraction

This section emphasizes abstract feature extraction with EfficientNet, prioritizing highly informative features from input images. The key advantage of EfficientNet's network scaling strategy is that it uses a compound coefficient ϕ to uniformly scale network width, depth, and resolution in a principled way. The approach enhances computational and parameter efficiency, enabling the extraction of high-quality features with reduced computational demands.

$$
\begin{aligned}
depth &: d = \alpha^{\phi} \\
width &: \omega = \beta^{\phi} \\
resolution &: r = \gamma^{\phi}
\end{aligned}
\tag{1}
$$

where α, β, and γ can be determined through a small-scale search. ϕ governs the allocation of additional resources for model expansion, while α, β, and γ dictate how these supplementary resources are distributed across network width, depth, and resolution. We have established the following constraints:

$$
\begin{aligned}
\text{s.t. } &\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2 \\
&\alpha \geq 1, \beta \geq 1, \gamma \geq 1
\end{aligned}
\tag{2}
$$

By increasing the model's depth, it becomes proficient at capturing more abstract features. Increasing the model's width helps in create diverse feature sets. Raising the input image's resolution aids in capturing finer details. The balance among these three aspects is crucial because it enables the model to utilize image features more effectively without compromising performance, thereby enriching the information extracted from images.

EfficientNet also employs a multi-scale feature extraction strategy by stacking convolutional layers with different

resolutions, effectively capturing both fine-grained details and abstract features within images. When dealing with images of varying resolutions, it excels at feature extraction. The choice of EfficientNet-B0 over other versions (B1-B7) is typically driven by a consideration of balancing performance and computational cost. B0 represents the most lightweight model within the EfficientNet series, making it suitable for executing image feature extraction tasks in scenarios with limited computational resources [13].

2.2 Feature Classification

This section is responsible for the analysis and categorization of the extracted features. High-dimensional feature maps are reduced to suitable low-dimensional representations through the utilization of connectivity blocks. The following are the specific functions of the internal levels of DNM [14]. It includes synapse, dendritic, menbrance, and soma layers.

Synapse layer: In the forward propagation phase, the input data initially passes through the synapse layer, where it undergoes operations including weight multiplication, bias subtraction, and ReLU activation. This layer simulates the synaptic functions of neurons, involving information transmission and nonlinear information processing. The output of the j-th dendritic branch for the i-th element of the input x_i can be defined as: $A_{ij} = ReLU(k \cdot (w_{ij} \cdot x_i - p_{ij}))$. This operation includes multiplication between the weights w_{ij} and the input x_i, followed by the subtraction of the threshold p_{ij}. Subsequently, the ReLU activation function is applied, which restricts negative values to zero, introducing non-linearity and aiding in the preservation of relevant signals while suppressing noise and irrelevant information. The learnable parameters, w_{ij}, and p_{ij}, are randomly initialized within the range of $(0, 1)$. The amplification of the resulting signal is regulated by k, which is initialized to 0.5.

Dendritic layer: It receives the output from the preceding synaptic layer and calculates the cumulative results for individual dendritic branches. This function replicates the behavior of neuronal dendrites, consolidating information from multiple synapses. The j-th dendritic branch combines N input signals A_{ij}, S_j is the output of dendritic layer, and the output can be expressed as $S_j = \sum_{i=1}^{N} A_{ij}$.

Membrane layer: It accumulates signals from all dendritic branches through a summation operation. It consolidates the outputs of M dendritic branches to produce a comprehensive representation denoted as $V = \sum_{j=1}^{M} S_j$. The variable V signifies the overall input integration from the dendritic layer.

Soma layer: It applies the ReLU activation function to the output of the membrane layer. The final output can be written as $O = ReLU(k_s \cdot (V - t_s))$. k_s and t_s are learnable parameters initializing in the range of $(0, 1)$.

These two components collaborate seamlessly, fully harnessing the feature extraction capacity of EfficientNet-B0 and the nonlinear information processing capabilities of DNM to bolster task performance such as image classifica-

Table 1: The experimental results of DEN and its peers.

Model	Accuracy(%)	Precision(%)	Recall(%)	F1(%)
MobileNet	84.4	89.1	81.8	85.0
Resnet50	82.6	81.1	84.4	82.1
Wideresnet50	83.9	91.1	87.4	88.7
Resnet152	79.1	82.4	77.8	79.5
Densenet161	84.1	89.0	87.6	88.8
DEN	**90.4**	**93.5**	**88.0**	**90.5**

Table 2: Parameter sensitivity analysis of DEN.

M	Accuracy(%)	Precision(%)	Recall(%)	F1(%)
2	88.7	90.8	87.6	88.9
4	89.7	90.5	89.6	89.8
6	**90.4**	**93.5**	88.0	**90.5**
8	88.6	86.7	90.8	88.3
10	89.1	87.7	**91.9**	89.5

tion. DEN effectively emulates the neural information processing mechanism, with the potential to enhance the precision in capturing and representing intricate image features.

3. Experiments and Results

The DR database from Kaggle was used, which includes 1020 diabetic retinopathy cases and 993 normal cases. The dataset was divided into a training set (70%) and a testing set (30%). We resize the DR dataset to a 512x512 resolution, which served as inputs for the DEN model. All experiments are conducted on a system running Ubuntu 18.04, equipped with an Nvidia RTX 3090 24GB GPU. Our implementation utilizes PyTorch 1.11 and Python 3.8, encompassing various neural networks, including DEN, MobileNet, ResNet50, DenseNet161, Wide-ResNet50, and ResNet152. Network parameters are optimized using the Adam optimizer with a learning rate of 0.001, and training is carried out for 100 epochs with a batch size of 32 for all experiments.

Table 1 summarizes the experiment results of all models. In the case of the DR problem, it is evident that DEN outperforms ResNet50 by 7.8% in Accuracy, 12.4% in Precision, 3.6% in Recall, and 8.4% in F1 score. It also outperforms Wide-ResNet50 by 6.5% in Accuracy, 2.4% in Precision, 0.6% in Recall, and 1.8% in F1 score. Furthermore, DEN surpasses ResNet152 by 11.3% in Accuracy, 11.1% in Precision, 10.2% in Recall, and 11% in F1 score.

ResNet50 typically exhibits better resistance to overfitting on small datasets. In contrast, Wide-ResNet50 benefits from its wider network structure, encompassing more convolutional layers and channels, thus enhancing the model's representational capacity. On the other hand, ResNet152, being deeper and more complex, is more susceptible to overfitting when dealing with smaller datasets. However, our employed DEN model effectively addresses this issue. It not only achieves a balanced arrangement in terms of network width, depth, and image resolution.

Table 3: Ablation Study on Diabetic Retinopathy.

Model	Accuracy(%)	Precision(%)	Recall(%)	F1(%)
Efficientnet-B2	87.2	85.1	89.8	87.1
Efficientnet-B2+DNM	89.2	88.8	**90.1**	89.3
Efficientnet-B0	88.4	88.8	88.4	88.3
DEN	**90.4**	**93.5**	88.0	**90.5**

3.1 The analysis of M

The specific hyper-parameter we primarily focus for DEN is the number of dendrites. The configuration of M parameter is essential to ensure the optimal performance of the DNM model in medical image tasks. The choice of the value for M is based on the task's characteristics. In scenarios characterized by limited data and relatively straightforward problems, a lower M value typically yields superior performance. Conversely, in the context of larger and more intricate tasks, a higher M value becomes a more suitable choice. Acquiring data for medical image problems is frequently challenging, resulting in a limited dataset, such as the DR fundus images in our study. Therefore, we conduct parameter discussion experiments on a smaller number of dendrites. Table 2 reveals that as M increased (M=2, 4, 6, 8, 10), each indicator exhibited a clear linear trend. Notably, we observe that the peak performance is achieved when M is set to 6.

3.2 Ablation Study

Table 3 presents the results of a series of ablation experiments performed to rigorously evaluate the effectiveness of our proposed methodology. The primary objective of these experiments is to comprehensively assess the impact of integrating the Efficientnet-B0, Efficientnet-B2, and DEN models on the classification performance of DR. The evaluation metrics employed include accuracy, precision, F1 score, and recall, which serve as quantitative measures for discerning the enhancements achieved through the incorporation of these components.

The performance improvement of the network is evident when comparing B0 with DEN, demonstrating the enhanced nonlinear capabilities of the DNM module. The comparison of EfficientNet-B0 and EfficientNet-B2 shows that a lightweight network is enough to solve the DR problem. Therefore, we chose EfficientNet-B0.

4. Conclusions

In our study, the proposed DEN model has demonstrated that when combining image feature extraction with nonlinear information processing, it achieved excellent results in the DR medical classification problem. It significantly enhances the performance of deep learning models. We evaluated models with varying numbers of dendrites on a dataset and compared them with other traditional classification network models. Our research results indicate that the DEN model outperforms the baseline model (EfficientNet-B0) and EfficientNet-B2 in terms of performance. Further-

more, compared to other network models, the DEN model exhibits stronger classification abilities on the DR dataset.

References

[1] M. Lotfy, J. Adeghate, H. Kalasz, J. Singh, and E. Adeghate, "Chronic complications of diabetes mellitus: A mini review," Current Diabetes Reviews, vol.13, no.1, pp.3–10, 2017.

[2] Y.S. Kanungo, B. Srinivasan, and S. Choudhary, "Detecting diabetic retinopathy using deep learning," 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), pp.801–804, IEEE, 2017.

[3] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.4510–4520, 2018.

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.770–778, 2016.

[5] G. Huang, Z. Liu, L. van der Maaten, and K.Q. Weinberger, "Densely connected convolutional networks," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.4700–4708, July 2017.

[6] M.E. Larkum, "Are dendrites conceptually useful?," Neuroscience, vol.489, pp.4–14, 2022. Dendritic contributions to biological and artificial computations.

[7] W.S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," The Bulletin of Mathematical Biophysics, vol.5, pp.115–133, 1943.

[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol.521, no.7553, pp.436–444, 2015.

[9] S. Gao, M. Zhou, Y. Wang, J. Cheng, H. Yachi, and J. Wang, "Dendritic neuron model with effective learning algorithms for classification, approximation, and prediction," IEEE Transactions on Neural Networks and Learning Systems, vol.30, no.2, pp.601–614, 2019.

[10] Z. Zhang, Z. Lei, M. Omura, H. Hasegawa, and S. Gao, "Dendritic learning-incorporated vision transformer for image recognition," IEEE/CAA Journal of Automatica Sinica, 2023. 10.1109/JAS.2023.123978.

[11] Z. Liu, Z. Zhang, M. Omura, R. Wang, and S. Gao, "Dendritic deep learning for medical segmentation," IEEE/CAA Journal of Automatica Sinica, 2023. 10.1109/JAS.2023.123813.

[12] S. Gao, M. Zhou, Z. Wang, D. Sugiyama, J. Cheng, J. Wang, and Y. Todo, "Fully complex-valued dendritic neuron model," IEEE Transactions on Neural Networks and Learning Systems, vol.34, no.4, pp.2105–2118, 2023.

[13] M. Tan and Q. Le, "EfficientNetV2: Smaller models and faster training," International Conference on Machine Learning, pp.10096–10106, PMLR, 2021.

[14] R.L. Wang, Z. Lei, Z. Zhang, and S. Gao, "Dendritic convolutional neural network," IEEJ Transactions on Electrical and Electronic Engineering, vol.17, no.2, pp.302–304, 2022.