ICIC Express Letters
Part B: Applications
Volume **13**, Number **9**, September **2022**
 ISSN 2185-2766
pp. **939–948**

EFFECTIVE DATA RESAMPLING AND META-LEARNING CONVOLUTIONAL NEURAL NETWORKS FOR DIABETIC RETINOPATHY RECOGNITION

Prem Enkvetchakul[1], Olarik Surinta[2] and Sangdaow Noppitak[1,*]

[1]Department of Information Technology
Faculty of Sciences
Buriram Rajabhat University
Naimuang Sub-District, Muang District, Buriram 31000, Thailand
prem.ev@bru.ac.th; *Corresponding author: sangdaow.np@bru.ac.th

[2]Multi-agent Intelligent Simulation Laboratory (MISL)
Department of Information Technology
Faculty of Informatics
Mahasarakham University
Khamriang Sub-District, Kantarawichai District, Mahasarakham 44150, Thailand
olarik.s@msu.ac.th

Received January 2022; accepted March 2022

Abstract. *Rapid diagnosis increases the chance of a patient being cured of symptoms. This applies especially to diabetic diseases where there is a high risk of diabetic retinopathy, which will lead to blindness if not treated promptly. Artificial intelligent techniques are proposed to diagnose diabetic retinopathy. In this paper, we recognize diabetic retinopathy from retinal images using meta-learning Convolutional Neural Networks (CNNs). Before training state-of-the-art CNNs, data resampling methods were proposed to select training and validation sets, and then the CNNs were trained on the selected training data. The simple data augmentation techniques were applied when training the CNNs to increase the training data pattern. We compared two ensemble learning methods: meta-learner and unweighted average, to show that the ensemble methods always performed better than when using a single CNN. The results showed that training the CNN model with the random data method outperformed other data resampling methods. However, data augmentation techniques did not present an outstanding result on diabetic retinopathy. In conclusion, the ensemble learning method using the meta-learner method resulted in the best accuracy when compared with unweighted average method. The proposed meta-learner CNNs achieved an accuracy of 86.32%.*
Keywords: Data resampling method, Lightweight convolutional neural networks, Ensemble learning method, Meta-learning method, Diabetic retinopathy recognition

1. **Introduction.** According to the World Health Organization (WHO), the world's population with diabetes is constantly on the rise [1]. Furthermore, according to data collected from a study in Henan Provincial People's Hospital in China, diabetic retinopathy was diagnosed in 580 (45.2%) out of 1,284 diabetic patients [2]. Also, other studies have shown that people with diabetes have a high risk of contracting diabetic retinopathy, a diabetic complication that can lead to significant vision loss or blindness. However, early detection of diabetic retinopathy greatly increases the chances of recovery while also helping to mitigate or even prevent the damage to a diagnosed diabetic's vision [3].

To make such early detection, though, ophthalmologists rely on examining patient retinal images to diagnose retinal diseases which can be a time-consuming process. Unfortunately, if a doctor is overworked with too many patients, a proper examination and diagnosis of a patient's retinal images can take too long to help the patient. Because

DOI: 10.24507/icicelb.13.09.939

the symptoms of different eye diseases can present very similarly, the images must be examined by a trained ophthalmologist in order to make a correct diagnosis of diabetic retinopathy [4]. However, due to advancements in both hardware technology and artificial intelligence, retinal identification [5] and eye automated diagnostic tools have been developed to detect diabetic retinopathy and its different stages by using retinal images [6]. Thus, the proper application of new diagnostic tools could help ophthalmologists reduce the time it takes to make a final positive diagnosis and begin a patient's treatment.

Related Work. Sarki et al. [4] proposed automated detection of mild and multi-class diabetic eye diseases using Convolutional Neural Network (CNN) architectures. Two CNN models, VGG16 and InceptionV3, were proposed to train retinal images and classify them into five classes. The results showed that the VGG16 outperformed InceptionV3. In addition, they evaluated the CNNs on the pre-trained and fine-tuning models. They found that the fine-tuning models of the VGG16 achieved the best accuracy on the diabetic eye diseases dataset.

Chen et al. [7] proposed a deep feature learning method for glaucoma detection. In their method, they initially proposed six CNN architectures. It contains five multilayer perceptron convolutional layers and one fully-connected layer. In the output layer, the softmax regression was applied. They also applied dropout with the value of 0.5 and data augmentation techniques: image translations and horizontal reflections. Secondly, the concatenated structure of CNNs was proposed to combine five CNNs. Their proposed method was evaluated on ORIGA and SCES datasets, the glaucoma fundus image datasets. The result showed that the concatenated structure of CNNs achieved an accuracy of 83.8% and 89.8% on the ORIGA and SCES datasets, respectively.

Nazir et al. [1] presented results of detecting diabetic eye disease from retinal images using deep learning. The first process involved preparing the dataset and feature extraction, while the second process involved improving a custom CenterNet model trained for ophthalmic classification. The DenseNet-100 was used as the feature extraction method. CenterNet was employed to localize and classify disease lesions. The recognition performance of this method was 97.93% on the APTOS-2019 dataset and 98.10% on the Indian Diabetic Retinopathy image Dataset (IDRiD).

For diabetic retinopathy recognition, Doshi et al. [8] proposed deep CNN architectures to recognize diabetic retinopathy from retinal images that contain five classes. They proposed a new CNN architecture that contains 14 layers and two layers of the fully-connected layer. The Leaky rectifier activation function was used in training. Hattiya et al. [9] proposed various CNN architectures to create the CNN models and recognize them as two classes. During image processing process, they transformed the retinal images from RGB color space into grayscale, HSV, L*a*b*, and YCbCr color spaces before sending them to train with CNNs. However, training the CNN using AlexNet architecture with RGB color space achieved the highest accuracy on 98.42% and 81.32% on training and test sets, respectively.

For the ensemble learning method, Qummar et al. [6] proposed ensemble CNNs that contain five CNNs to classify diabetic retinopathy into five classes: ResNet50, Inception-V3, Xception, DenseNet121, and DenseNet169. Mahmoud and Yaroshchak [10] proposed bagging ensemble learning that combines the output from three CNN architectures and uses a weighted average method to calculate the output. The random training data with replacement method was used in the bagging process. Hence, training data in set one can appear again in other sets. In their experiments, the retinal images were divided into four classes. The results showed that the bagging ensemble learning method increased accuracy compared with single CNN.

Contribution. This paper presents meta-learning CNNs architecture to recognize diabetic retinopathy from the retinal images. First, we compared four data resampling methods: random data, fixed data, 5K-fold, and bootstrap. Second, the resampling data from the first step is given to create the robust CNN models with lightweight CNNs: MobileNetV2, EfficientNetB1, and NASNetMobile. In this step, we also performed the simple data augmentation techniques: rotation, zoom, vertical flip, and horizontal flip, that augment the images without distortion. Finally, ensemble learning methods, including meta-learner and unweighted average, were compared. In this meta-learner, the various machine learning methods (logistic regression: LR, support vector machine: SVM, K-nearest neighbors: KNN, and random forest: RF) were proposed to learn the probability values computed from the lightweight CNNs and classify the retinal images. While the unweighted average method averages the probability values, the maximum probability value was selected as the predicted label.

Paper Outline. The detail of the meta-learning CNNs architectures is explained in Section 2. The experimental results and conclusion are presented in Section 3 and Section 4.

2. **Meta-Learning Convolutional Neural Networks Architecture.** In this section, we present the meta-learning CNNs to recognize retinal images on the Diabetic Retinopathy (DR) dataset. The meta-learning CNNs architecture is designed to learn from the probabilities of state-of-the-art CNNs. It consists of two levels: CNN-learner and meta-learner. The CNN-learner trains on the training data, which is the retinal images. Further, the meta-learner trains on the probability values computed by CNN learners. The framework of the proposed method is shown in Figure 1.

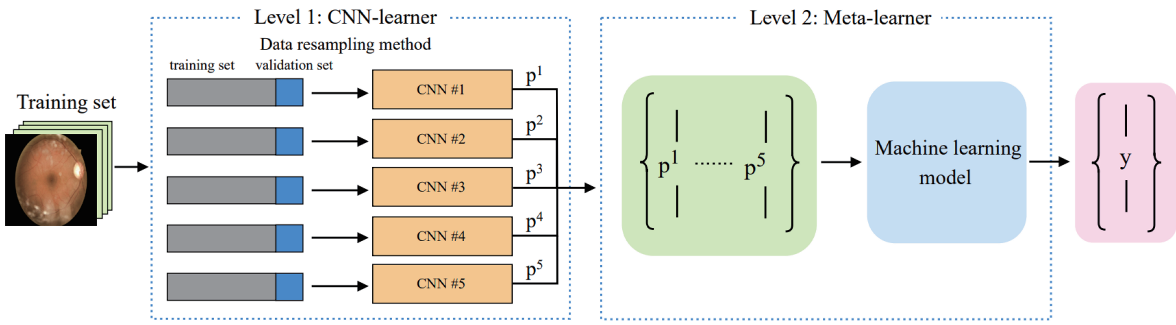


FIGURE 1. Illustration of the meta-learning CNNs architecture

2.1. **Data resampling methods.** The outstanding representation data will increase the robustness of the model for the recognition tasks. In this research, we mainly focus on selecting training and validation sets, while comparing the effectiveness of four various data resampling methods: random data, fixed data, 5K-fold, and bootstrap. The data resampling methods contain four methods as follows.

Random data. Under the random data method, we randomly allocated 80% of the data to the training set and 20% to the validation set. We performed five separate random data allocations to create five distinct training and validation sets, as shown in Figure 2(a).

Fixed data. The data was randomly selected only one time. This method selected 80% for training and 20% for validation sets. Further, CNN models were trained with the same training set, as shown in Figure 2(b).

5K-fold. The idea of splitting the data into five folds, called the 5K-fold method, was applied. For the 5K-fold method, each fold contains data of around 20%. The training set contains 4-fold, while the validation set contains only 1-fold. The validation set was selected from fold five until fold one in each training [11], as shown in Figure 2(c).

Bootstrap. This method allocated the training of data in accordance with the replacement method [12], which also allows for the use of duplicate images to be allocated to

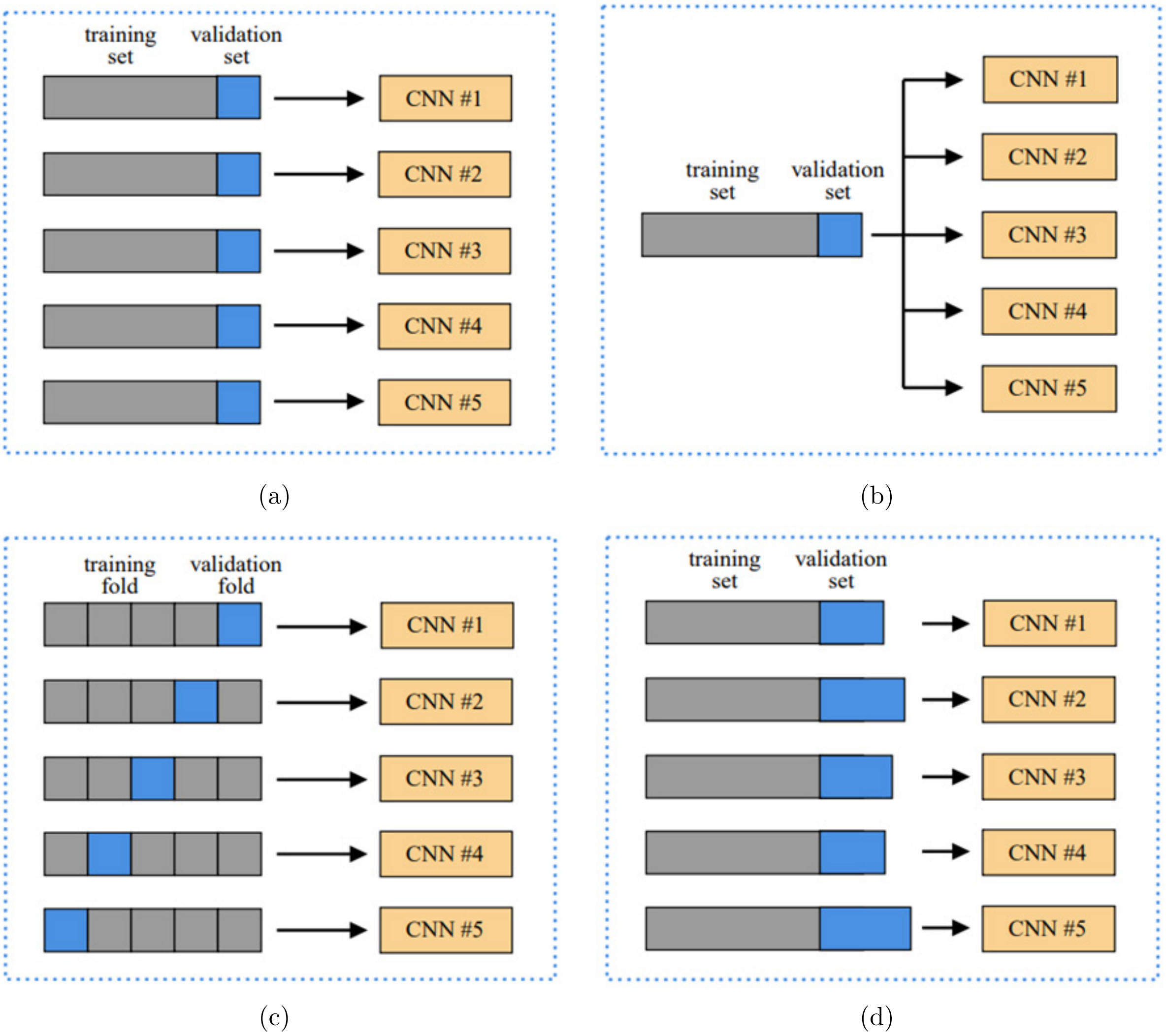


FIGURE 2. Illustration of the data resampling methods: (a) Random data, (b) fixed data, (c) 5K-fold, and (d) Bootstrap

sets. In our experiment, the training dataset allocated equal amounts of data to the fixed data method, random data method, and 5K-fold method. However, the amount of data allocated to the validation set was dependent on the images not allocated to the training dataset. As a result, the amount of data in the validation data sets could vary with each separate random allocation, as shown in Figure 2(d).

2.2. **Convolutional neural networks.**

MobileNetV2. MobileNetV2 was proposed by Sandler et al. [13]. It is a lightweight CNN that includes inverted residual structure and linear bottlenecks. The bottlenecks block contains three transforming layers: convolution 2D, depthwise separable convolution, and linear convolution 2D. Further, the shortcut connections are used to link between two bottlenecks. When comparing the MobileNetV1 and V2, the MobileNetV2 has lower parameters and computation time than MobileNetV1.

EfficientNetB1. EfficientNetB1 was proposed by Tan and Le [14]. It has eight models, including EfficientNetB0-EfficientNetB7. The model EfficientNetB0 is the smallest size, while the EfficientNetB7 is the largest. The model of the EfficientNet is scaling by considering four factors: width, depth, resolution, and compound methods, to balance the network. The EfficientNetB0 was designed to be similar to MnasNet, but scaling factors were applied that directly affected the size of the model.

NASNetMobile. Zoph et al. [15] proposed a search architecture, called Neural Architecture Search Network (NASNet). It contains two types of cells: normal and reduction cells. First, the normal cells were built by searching several convolution operations using a recurrent neural network and reinforcement learning. Second, the spatial dimension of the feature maps from the normal cell was reduced and was called the reduction cell. Finally, the normal cell and reduction cell were stacked together. Furthermore, the normal cell could compute more than one time. The NASNetMobile has few cells repeats and filters in the network.

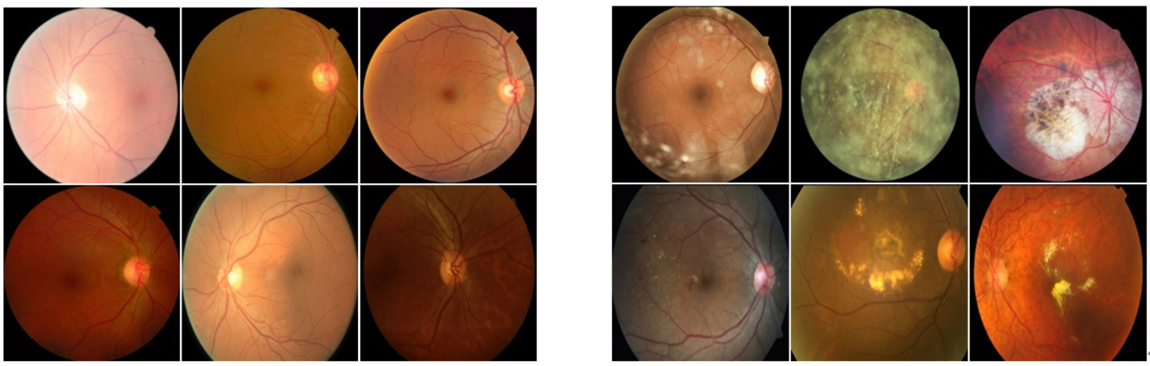
2.3. **Ensemble learning methods.** We examined ensemble learning methods to recognize diabetic retinopathy from the retinal images, including meta-learning and unweighted average. A short description of the ensemble methods is as follows.

Meta-learner. The meta-learner [16] was proposed to learn the probabilities that were the output of state-of-the-art CNN-learners. It was the second level of our proposed architecture. For the meta-learner, we investigated four machine learning techniques consisting of LR, SVM, KNN, and RF to learn the output probabilities of the CNNs and create robust models.

Unweighted average. The unweighted average is the standard method that is computed by averaging the output probabilities of the classifiers [17]. In this paper, we averaged the output probabilities of the CNNs that were calculated using the softmax function. Then, the highest probability value was selected as the final prediction.

3. **Experimental Setting and Results.** In this research, all the retinal images were transformed into $224 \times 224 \times 3$ pixels resolution according to the three CNN architectures, EfficientNetB1, MobileNetV2, and NASNetMobile architectures. The pre-trained models were used and we trained the CNN models with these hyperparameters: epoch = 100, batch size = 16, learning rate = 0.01, and optimizer = Stochastic Gradient Descent (SGD). The TensorFlow library was used as the deep learning platform. All experiments were run on the Linux operating system using Intel(R) Core-15, 2320 CPU @ 3.00GHz, 16GB RAM, and GPU GeForce GTX 1060Ti.

3.1. **Diabetic retinopathy dataset.** Three benchmark DR datasets were published on the Kaggle website: APTOS 2019 blindness detection, diabetic retinopathy, and diabetic retinopathy detection. In these datasets, the retinal images with labels were verified by the experts of the Aravind eye hospital [8]. We randomly selected 23,510 retinal images and divided them into training and test sets with 80 : 20 ratios. The retinal images contain 11,444 non-diabetic retinopathy and 12,066 diabetic retinopathy images, as shown in Figure 3.



(a) (b)

FIGURE 3. Examples of the (a) non-diabetic retinopathy and (b) diabetic retinopathy images

3.2. **Experiments on lightweight CNNs.** In this experiment, three lightweight CNN models, MobileNetV2, EfficientNetB1, and NASNetMobile, were examined. We also applied basic Data Augmentation (DA) techniques: rotation, zoom, vertical flip, and horizontal flip. The DR dataset was divided into training and test sets with 18,808 and 4,702 images, respectively. The training set was split into training and validation sets using four data resampling methods: random data, fixed data, 5K-fold, and bootstrap. Five experiments were used to evaluate the CNN models, and the average accuracy and standard deviation were reported.

The experimental results of the data resampling showed that the random data method outperformed fixed data, 5K-fold, and bootstrap methods on the test set. It can be seen from Table 1 that all data resampling methods achieved higher test accuracy than the average accuracy on the validation set. As a result, the EfficientNetB1, when choosing the training set with a random data method, achieved the highest accuracy with 84.05%. The worst performances were obtained when training the CNN model with applied data augmentation techniques but did give better results with the validation set.

Table 1. Performance of the CNNs and data resampling methods on the DR dataset

CNN architectures	**DA**	**Random data**		**Fixed data**		**5K-fold**		**Bootstrap**	
		Validation (%)	**Test** (%)	**Validation** (%)	**Test** (%)	**Validation** (%)	**Test** (%)	**Validation** (%)	**Test** (%)
MobileNetV2	No	75.21±0.78	**84.77**	75.16±0.35	82.41	75.43±0.69	84.13	73.55±0.32	82.35
	Yes	78.98±0.57	82.45	79.46±0.48	80.41	79.28±0.45	82.26	77.68±0.33	81.01
EfficientNetB1	No	76.49±1.17	**84.05**	76.10±0.50	83.24	76.32±0.48	83.60	75.18±0.17	81.62
	Yes	78.97±1.09	82.18	78.01±0.24	82.41	78.18±0.39	82.88	76.87±1.14	80.92
NASNetMobile	No	81.55±1.22	83.33	73.46±1.48	81.05	73.36±0.18	83.05	72.65±0.46	81.41
	Yes	77.03±1.77	**83.59**	75.51±0.48	81.71	75.28±0.89	82.18	73.99±0.64	81.20

3.3. **Experiments on ensemble learning methods.** We examined the performance of ensemble learning using an unweighted average method on the test set of the DR dataset. First, we discovered the best number of ensemble CNN models by combining 1, 2, 3, 4, and 5 CNN models. Second, we trained the CNN models with and without applying data augmentation techniques to confirming that the data augmentation techniques do not impact the recognition performance.

In the data augmentation experiment, four data augmentation techniques were proposed: horizontal flip, vertical flip, rotation, and zoom. The random rotation value between 0-90° and zoom in and out with a random value of $[1 - 0.2, 1 + 0.2]$ were applied to the retinal images.

In the first experiment, the accuracies of different numbers of ensemble CNN models obtained in the first experiment are shown in Figure 4. It can be seen that size of the ensemble CNNs affected the accuracy of the result. We found that using five ensemble CNN models achieved the highest result when training with different data resampling methods.

In the second experiment, we selected only five ensemble models according to the results, as shown in Table 2. The results showed that the ensemble CNNs with MobileNetV2, when resampling data using random data, outperformed other N architectures and other data resampling data methods. An accuracy of 86.01% was achieved. Also, comparing the ensemble CNNs with the single CNN confirmed that the ensemble CNNs present the best performance with approximately 1.2%.

3.4. **Experiments on CNNs with the meta-learning method.** To examine the meta-learning method, we first received the probability values of the retinal images that were computed using the fine-tuned CNN models. Second, the probability values were

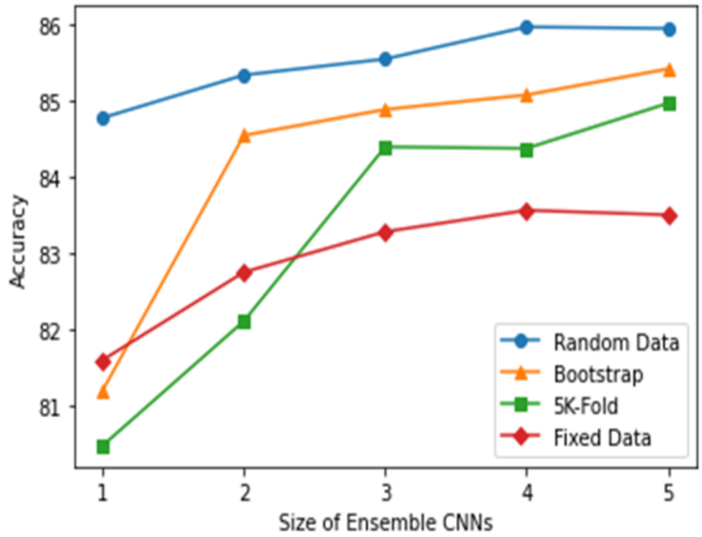


Figure 4. The performance of the ensemble CNNs when trained with various data resampling methods

Table 2. The performance of diabetic retinopathy recognition experiments on three CNN architectures with different ensemble learning methods and data augmentation techniques

CNN architectures	**DA**	**Accuracy (%) of data resampling methods**			
		Random data	**Fixed data**	**5K-fold**	**Bootstrap**
MobileNetV2	No	**86.01**	83.50	84.96	85.41
	Yes	82.50	81.34	82.80	82.77
EfficientNetB1	No	85.79	84.26	85.09	83.54
	Yes	82.34	82.65	84.75	83.13
NASNetMobile	No	85.28	82.96	84.69	83.86
	Yes	83.18	82.20	84.01	83.50

transferred to the machine learning techniques to create the robust models. We also adjusted the machine learning models with the following parameters. SVM: C = 1, and gamma = 0.1, kernel = RBF, KNN: K = 19, distance value = Euclidean, and weight = uniform, RF: estimators = 800, max depth = 30, min samples leaf = 4, min samples split = 10, min features = auto, and bootstrap = true.

Table 3 shows the results of experiments with the meta-learner methods. We found that when applying the data resampling method using random data, the methods consistently outperformed other data resampling methods. In this experiment, the highest accuracy was obtained when the retinal images were trained with MobileNetV2 and then transferred the probability values to the LR method to create the meta-learner model. It achieved an accuracy of 86.32% on the DR dataset. Consequently, the accuracy of the meta-learner method was slightly higher than the ensemble CNN method when evaluated on the DR dataset.

The confusion matrices of the unweighted average method used MobileNetV2 as a CNN architecture and meta-learning method combined with logistic regression on the diabetic retinopathy dataset, as shown in Figure 5. The meta-learning method technique has the advantage of decreasing classification errors. Recognition performance and incorrect classification from DR to non-DR class were reduced from 248 to 214 instances.

We also considered the testing time of the ensemble CNN and meta-learning methods. The testing time showed that the NASNetMobile spent more computation time on the testing set of the DR dataset with both ensemble CNN and meta-learning methods. It

Table 3. The accuracy performance of the meta-learning method experiments on four machine learning techniques and three CNN architectures

Machine learning techniques	CNN architectures	Accuracy (%) of data resampling methods			
		Random data	Fixed data	5K-fold	Bootstrap
Logistic regression	MobileNetV2	**86.32**	83.16	84.56	85.09
	EfficientNetB1	85.52	84.24	84.90	83.39
	NASNetMobile	84.81	82.82	84.88	83.60
Support vector machine	MobileNetV2	86.13	83.26	84.18	85.07
	EfficientNetB1	85.60	84.64	85.28	83.33
	NASNetMobile	84.92	81.94	84.62	83.69
K-nearest neighbors	MobileNetV2	86.03	83.26	84.45	85.16
	EfficientNetB1	85.69	84.54	85.30	84.03
	NASNetMobile	84.86	81.43	84.64	83.86
Random forest	MobileNetV2	85.58	82.69	84.45	85.84
	EfficientNetB1	85.73	84.18	85.03	84.67
	NASNetMobile	84.47	82.58	85.05	84.24

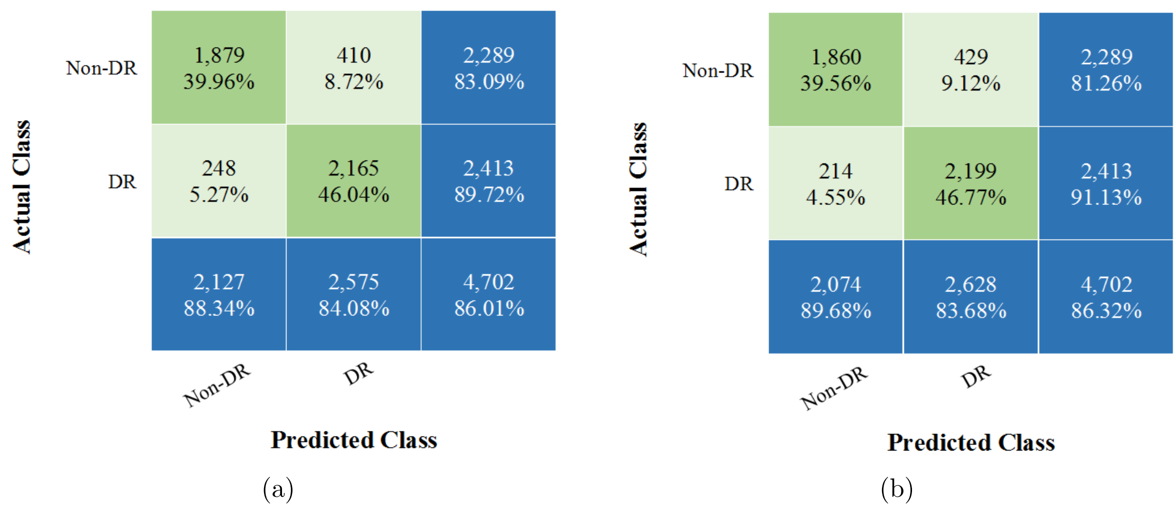


Figure 5. Illustrated confusion matrices of (a) the unweighted average method using MobileNetV2 architecture as a CNN architecture and (b) meta-learning method combined with logistic regression on the diabetic retinopathy dataset

spent around four milliseconds per image on the ensemble CNNs and meta-learning methods. As a result, MobileNetV2 took only approximately 13 milliseconds. Compared to the training time of single CNN, the MobileNetV2 was trained faster than EfficientNetB1 and NASNetMobile. MobileNetV2 spent 5 hours and 42 minutes when EfficientNetB1 and NASNetMobile trained for 13 hours 34 minutes and 17 hours 31 minutes, respectively. Consequently, the meta-learning and ensemble CNNs methods computed almost the same testing time.

4. **Conclusions.** This paper proposed the meta-learning CNNs framework to recognize diabetic retinopathy images, which were recognized as two classes: diabetic retinopathy and non-diabetic retinopathy. We mainly experimented with data resampling methods, including random data, fixed data, 5K-fold, and bootstrap.

For the experiments, we first proposed three state-of-the-art lightweight CNN architectures: MobileNetV2, EfficientNetB1, and NASNetMobile, to create a robust CNN model by training on the retinal images. Second, we created the ensemble CNNs learning with an unweighted average method. In the ensemble CNNs, five CNN models were investigated. Finally, the meta-learning method was proposed, which is the stacking of CNN models and machine learning methods. We compared the experiments with four machine learning methods: LR, SVM, KNN, and RF. We especially experimented with various data resampling methods, including random data, fixed data, 5K-fold, and bootstrap. We discovered that the random data method consistently outperformed other data resampling methods.

In addition, the experimental results obtained showed that the lightweight MobileNetV2 outperformed the EfficientNetB1 and NASNetMobile in terms of accuracy and testing time. The meta-learning method, which stacks the MobileNetV2 and LR methods, slightly improved accuracy compared with the ensemble CNNs. Both methods achieved around 86% accuracy on the diabetic retinopathy dataset.

In future work, we aim to create the fusion CNNs and extract the deep features using both CNN and Recurrent Neural Networks (RNNs).

REFERENCES

[1] T. Nazir, M. Nawaz, J. Rashid, R. Mahum, M. Masood, A. Mehmood, F. Ali, J. Kim, H.-Y. Kwon and A. Hussain, Detection of diabetic eye disease from retinal images using a deep learning based CenterNet model, *Sensors*, vol.21, no.16, DOI: 10.3390/s21165283, 2021.

[2] X. Yao, X. Pei, Y. Yang, H. Zhang, M. Xia, R. Huang, Y. Wang and Z. Li, Distribution of diabetic retinopathy in diabetes mellitus patients and its association rules with other eye diseases, *Sci. Rep.*, vol.11, no.1, pp.1-10, 2021.

[3] T. Nazir, A. Irtaza, A. Javed, H. Malik, D. Hussain and R. A. Naqvi, Retinal image analysis for diabetes-based eye disease detection using deep learning, *Appl. Sci.*, vol.10, no.18, DOI: 10.3390/app 10186185, 2020.

[4] R. Sarki, K. Ahmed, H. Wang and Y. Zhang, Automated detection of mild and multi-class diabetic eye diseases using deep learning, *Heal. Inf. Sci. Syst.*, vol.8, no.1, pp.1-9, 2020.

[5] I. Kanjanasurat, B. Purahong, H. Aoyama, C. Benjangkaprasert and C. Pintavirooj, Personal identification using a delaunay triangle and optic disc retinal vascular pattern, *International Journal of Innovative Computing, Information and Control*, vol.16, no.3, pp.879-897, 2020.

[6] S. Qummar, F. Khan, S. Shah, A Khan, S. Shamshirband, Z. Rehman, I. Khan and W. Jadoon, A deep learning ensemble approach for diabetic retinopathy detection, *IEEE Access*, vol.7, pp.150530-150539, 2019.

[7] X. Chen, Y. Xu, S. Yan, D. W. K. Wong, T. Y. Wong and J. Liu, Automatic feature learning for glaucoma detection based on deep learning, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Lecture Notes in Computer Science*, N. Navab, J. Hornegger, W. Wells and A. Frangi (eds.), Cham, Springer, DOI: 10.1007/978-3-319-24574-4_80, 2015.

[8] D. Doshi, A. Shenoy, D. Sidhpura and P. Gharpure, Diabetic retinopathy detection using deep convolutional neural networks, *International Conference on Computing, Analytics and Security Trends (CAST)*, pp.261-266, 2017.

[9] T. Hattiya, K. Dittakan and S. Musikasuwan, Diabetic retinopathy detection using convolutional neural network: A comparative study on different architectures, *Eng. Access*, vol.7, no.1, pp.50-60, 2021.

[10] S. Mahmoud and S. Yaroshchak, Bagging of convolutional neural networks for diagnostic of eye diseases, *The 4th International Conference on Computational Linguistics and Intelligent Systems (COLINS)*, pp.715-729, 2020.

[11] E. G. Adagbasa, S. A. Adelabu and T. W. Okello, Application of deep learning with stratified K-fold for vegetation species discrimation in a protected mountainous region using Sentinel-2 image, *Geocarto Int.*, pp.1-21, 2019.

[12] A. R. Hassan and M. I. H. Bhuiyan, Computer-aided sleep staging using complete ensemble empirical mode decomposition with adaptive noise and bootstrap aggregating, *Biomed. Signal Process. Control*, vol.24, pp.1-10, 2016.

[13] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. C. Chen, MobileNetV2: Inverted residuals and linear bottlenecks, *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.4510-4520, 2018.

[14] M. Tan and Q. V. Le, EfficientNet: Rethinking model scaling for convolutional neural networks, *The 36th International Conference on Machine Learning (ICML)*, pp.10691-10700, 2019.

[15] B. Zoph, V. Vasudevan, J. Shlens and Q. V. Le, Learning transferable architectures for scalable image recognition, *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.8697-8710, 2018.
[16] M. Gour and S. Jain, Stacked convolutional neural network for diagnosis of COVID-19 disease from x-ray images, *arXiv.org*, arXiv: 2006.13817, 2020.
[17] T. Chompookham and O. Surinta, Ensemble methods with deep convolutional neural networks for plant leaf recognition, *ICIC Express Letters*, vol.15, no.6, pp.553-565, 2021.