Application of Deep Convolutional Neural Networks VGG-16 and GoogLeNet for Level Diabetic Retinopathy Detection

Article · January 2022

DOI: 10.1007/978-3-030-89880-9_5

4 authors, including:

Komgrit Leksakul
Chiang Mai University

Pranprach Wattana
Chiang Mai University

Poti Chaopaisarn
Chiang Mai University

Application of Deep Convolutional Neural Networks VGG-16 and GoogLeNet for Level Diabetic Retinopathy Detection

Chaichana Suedumrong[1,2], Komgrit Leksakul[2(✉)], Pranprach Wattana[2], and Poti Chaopaisarn[2]

[1] Graduate Program/ Ph.D.'s Degree Program in Industrial Engineering, Department of Industrial Engineering, Faculty of Engineering, Chiang Mai University, Chiang Mai, Thailand
[2] Department of Industrial Engineering, Faculty of Engineering, Chiang Mai University, Chiang Mai, Thailand
komgrit@eng.cmu.ac.th

Abstract. Diabetic retinopathy (DR) is a diabetes complication that damages the retina. This type of medical condition affects up to 80% of patients with diabetes for 10 or more years. The expertise and equipment required are often lacking in areas where diabetic retinopathy detection is most needed. Most of the work in the field of diabetic retinopathy has been based on disease detection or manual extraction of features. Thus, this research aims at automatic diagnosis of the disease in its different stages using deep learning neural network approach. This paper presents the design and implementation of Graphic Processing Unit (hereby GPU) accelerated deep convolutional neural networks to automatically diagnose and thereby classify high-resolution retinal images into five stages of the disease based on its severity. The accuracy of the single model convolutional neural networks presented in this paper is 71.65% from VGG-16.

Keywords: Diabetic retinopathy · Deep learning · Convolutional neural networks · VGG-16 · GoogLeNet

1 Introduction

Diabetic retinopathy (DR) is a common blindness cause by diabetes complication which is mostly found in the United States. From the epidemic of obesity and diabetes in the USA and around the world, serious and frequent complications of diabetes have been developed as a significant public health problem, especially among minorities [1]. Diabetic retinopathy is diabetes-related damage to the retina, affecting up to 80% of people with diabetes for 10 years or more. Most diabetic retinopathy-related jobs depend on the identification of disease or manual feature extraction [2].

DR It is a diabetes complication which causes vision loss, resulting in one of the leading causes of blindness worldwide. According to statistics from the World Health Organization (WHO) in 2012, DR is the second leading cause of vision loss after cataracts.

K. Arai (Ed.): FTC 2021, LNNS 0000002, pp. 1–10, 2022.
https://doi.org/10.1007/978-3-030-89880-9_5

Blindness triggered by DR is caused by the destruction of small blood vessels over a long period of time. On average, DR occurs in approximately 20% of people with diabetes. Approximately 80% of patients with diabetes have been diagnosed with DR with approximately 10% of severe vision loss and 2% of blindness worldwide. Nearly 93 million people with DR, 28 million people have severe DR, which runs a high risk of vision loss [3]. The barriers to DR patients are increasing because of the high number of visits and the cost of testing [4].

Research has [5] shown that if switching to semi-automatic screening from healthcare professionals, costs can be reduced by 20%. In Thailand, the number of people with diabetes increases every year and it is estimated that at least 3 million people have diabetes. Diabetes is one of the top ten leading causes of death with death rate estimated at 10.8 per 100,000 people. The main cause of the complications arises in various systems and human organs, including blood vessels, heart, eyes, kidneys and feet. In Thailand, a considerable amount of money was invested in diabetes care. In 2008, it was found that the cost of treating a person with diabetes is approximately 28,207 baht per year [6]. Early-stage examination of DR is an extremely important task for people with diabetes as it would help in mitigating further complications which often leads to organ failures or loss of life.

To screen for DR, patients will be instilled with retinal dilation drugs. Pictures of the retina are taken with a retinal camera which provides a preliminary result. Then the retinal imaging results were sent to the ophthalmologist for further diagnosis. Based on the data from the Institute of Medical Technology Research and Evaluation, Department of Medicine, 2012, it was found that there are 328 ophthalmologists in Thailand, but there are about 5 million people with diabetes in Thailand, making screening for treatment possible slow. Furthermore, ophthalmologists are highly concentrated in large cities, making screening for DR a challenge nation-wide. To resolve this issue, automate DR checking is proposed to reduce the burden and reduce costs for both the hospital and the patient in terms of examination and service period [7–10].

VGG16 is a convolutional neural network (CNN) model presented by Simonyan and Zisserman in their research titled, "Very Deep Convolutional Networks for Large-Scale Image Recognition". The model is a widely recognized as it achieves 92.7% with a top-5 test accuracy in ImageNet, which is a dataset of over 14 million images belonging to 1000 classes [11]. VGG-16 has a high accuracy classification images capability in lung cytological images [12]. In classifying the eye status under near and open, GoogLeNet achieves non-misrecognition. When judging left or right faces, machine learning achieves a higher recognition level than deep learning. The results give us an indication for designing a higher accuracy recognition of the eye status to detect driver drowsiness [13]. This paper introduces GPU acceleration in deep convolutional neural networks (CNN) development and implementation VGG-16 and GoogLeNet to automatically diagnose and classify high-resolution retina images by intensity into five stages of the disease and examined their efficacy by comparing the accuracy rates from EyePACS data.

Given the shortage of ophthalmologists and the large number of diabetic patients who need diabetic retinopathy exams, we can alleviate the shortage of ophthalmologists with the advanced technology with tools to detect the level of diabetic retinopathy.

2 Convolutional Neural Network (CNN)

Convolutional Neural Network (hereby CNN) has the advantage of being able to extract features for ease of further classification. CNN is suitable for learning through pictures recognition. The neural network, a multilayer perceptron (multilayer perceptron), was first proposed by Yann LeCun [12, 13]. Back-Propagation returns the error value to adjust the error weight to the proper value. This is a highly popular neural network model as it allows faster workflow with high accuracy when compared with previous methods. Using weight adjustment to reduce the error function with the Back-Propagation enables the feature to be automatically extracted. Convolutional Filter generated by the weight value of Feature Maps by the synthetic neural network, the structure is shown in Fig. 1.

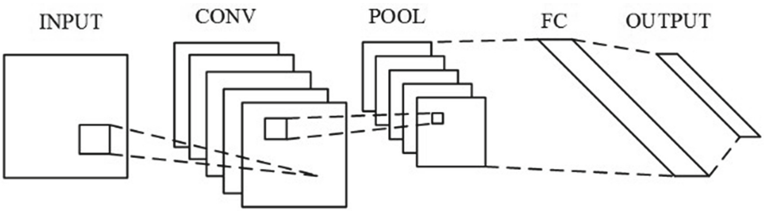

Fig. 1. Convolutional neural network architecture

According to Fig. 2, an example of converting between a 4×4 Pixel input and a 2×2 filter, results in a 3×3 Pixel output. Referring to the example output, the filter is coefficient and converts into input and output. The filter is then moved to next input data and the process is repeated until the size of input is complete. The output will be gradually reduced as this process slowly loses the edge of the image.

The working example of LeCun's LeNet-1 [14, 15] Neural Network, the first developed and improved neural network. LeCun used this neural network for handwritten USA letter-based numerical recognition. The images used were handwritten numerical images, and 10 numbers were classified: 0 to 10, the structure is shown in Fig. 3.

Quadratic Weighted Kappa Score [16] is a statistical measure measuring the level of consensus and inconsistency in between two commentators. The maximum possible value is 1, meaning that all parties agree, and 0 means none of the parties agree. $K, W_{i,j}, E_{i,j}, r_i$, and c_j are calculated by using Equation (1) - (5) where $W_{i,j}$ is the weighted matrix, $O_{i,j}$ is the multi class confusion matrix and $E_{i,j}$ being the expected matrix.

$$K = 1 - \frac{\Sigma_{ij} W_{ij} O_{ij}}{\Sigma_{ij} W_{ij} E_{ij}} \tag{1}$$

$$W_{i,j} = \frac{(i-j)^2}{(N-1)^2} \tag{2}$$

$$E_{ij} = n \times P\big(Y = i \, and \, \hat{Y} = j\big) = n \times P(Y = i) \times P\big(\hat{Y} = j\big) = n \times \frac{r_i}{n} \times \frac{c_j}{n} \tag{3}$$

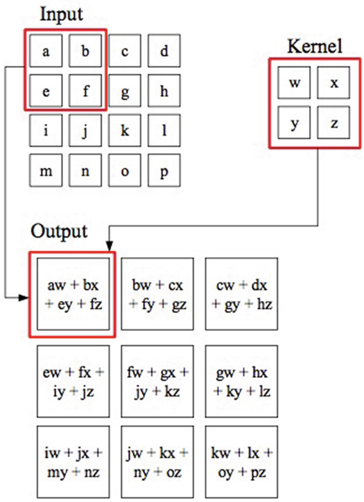

Fig. 2. Shows an example of convolution.

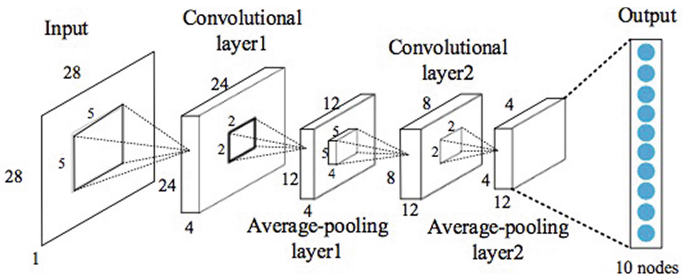

Fig. 3. Shows an example of LeNet-1

$$r_i = \Sigma_{i=i} O_{ij} \tag{4}$$

$$c_i = \Sigma_{j=j} O_{ij} \tag{5}$$

Compare test results from Test DataSet using Quadratic Weighted Kappa Score and accuracy measurements.

3 DATA Training and Validation

The purpose of this paper is to use the CNN approach with model learning of GoogLeNet and VGG-16 to create a new CNN with higher efficiency of diabetic retinopathy detection.

There are three important elements in this analysis. First, experimental data are prepared to ensure data validity with minimum error. Secondly, data of the Convolutional Neural Networks are then trained. Finally, the experimental results are then summarized and compared.

Data sets from EyePACS were used which were originally obtained from three Indian hospitals. Data is an 88,702 digital retinopathy photograph shown hereunder.

1. Screening damaged pictures as shown in the example in Fig. 4.
2. The data were screened through splitting all five data levels to have a comparable amount of data without overfitting and then the data were split into a training set.
3. Image were prepared by reducing the size of the image to 512 × 512 pixels and to reduce unwanted data, the black background was reduced as much as possible. After that, the data were duplicated into two sets, the first set are changed to Green Color Channel and the other are in RGB color as shown in the example in Fig. 5.

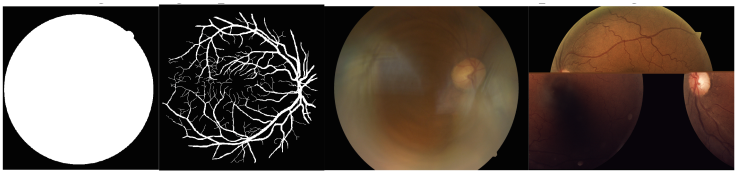

Fig. 4. Screening damaged pictures

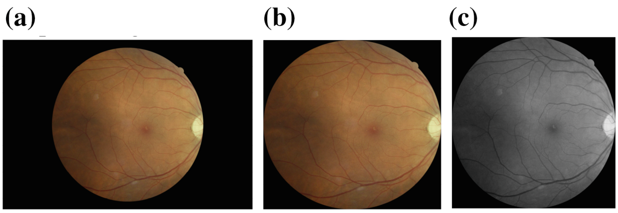

Fig. 5. a. Images before data preparation. b. Images after cropping off the black background. c. Images when changing to green color channel

To create a new CNN architecture with trainers, transfer learning method is used from GoogLeNet and VGG-16 models. By cutting off the fully connected layer and teaching the data by adding the Dense class to replace it (Table 1). After constructing the neural network, some fine-tuning is applied to the models.

Table 1. Neural network model convolution using transfer learning.

Layer	Dimension (pixel)	Layer	Dimension (pixel)
Input (1)	3 × 512 × 512	Dropout 6 (4)	400
Model (2)	GoogLeNet, VGG-16	Dense2 (5)	200
Dense 1 (3)	400	Output (6)	5

The experimental data is a set of fundus camera photo data, which is open-sourced data from EyePACs. The total data set is 88,702 images, divided into 35,126 of Training Set and 53,576 of Test Set images. The result of the Training Set has shown that a

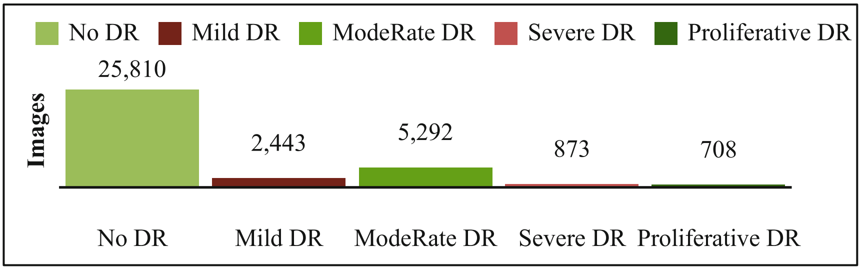

Fig. 6. Data of EyePACs.

considerate number of images results in no DR or Moderate DR of 25,810 out of 35,126 and 5,292 out of 35,126, respectively (Fig. 6).

Later, screening the number of images is conducted to prevent overfitting and increasing the number of images at other levels is done by creating a new image. Light adjustments of the image are performed to cut out the black sections to reduce the amount of unnecessary information. The total number of new data images is 77,362 presented in Fig. 7 and the division of the training data format Training, data segmentation training and inspection set are shown in Fig. 8.

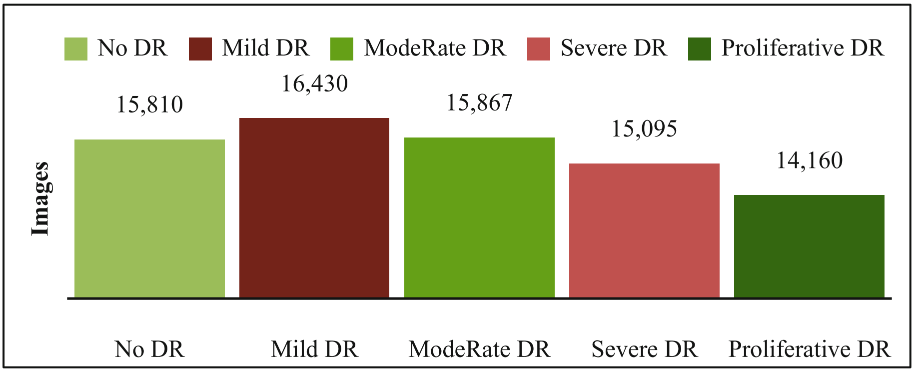

Fig. 7. Number of data images after creating new images.

Convolutional Neural Networks Creation and Training

CNN functions use deep learning tools such as, Keras with Tensorflow as a backend to evaluate the diabetic retinopathy process. The experiment uses 512 $\times$ 512 pixels as inputs. Ultimately, the output is a 5-level number of symptoms that are represented by a range of symptoms in ranging from level 0 to 4, as follows:

1) Inputting 100 rounds of training sessions of 4,000 images from the Training Set.
2) Evaluating the loss value of each round using images from the validation set.
3) Using Test Set results, comparing to the neural network model with Quadratic Weighted Kappa Score and Confusion Matrix.

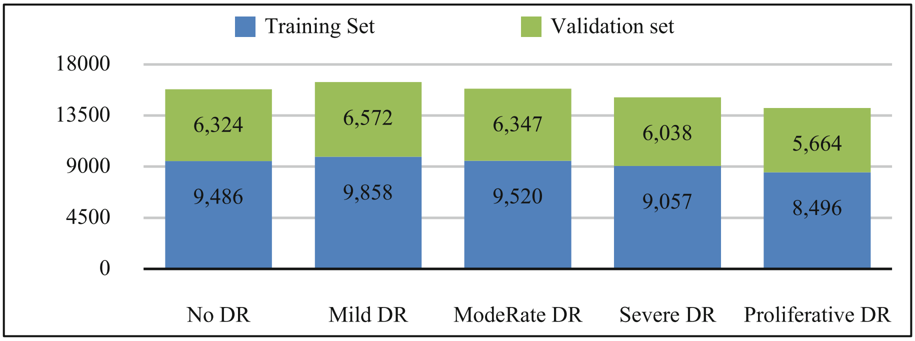

Fig. 8. Shows the number of data images in the training set and the validation set

Transfer Learning GoogLeNet and VGGNet-16 Models

For the Transfer Learning method, experimentation with the weight and structure from other databases are conducted and the output layer is changed to classify the disease level. An overview of the working structure is shown in Fig. 9.

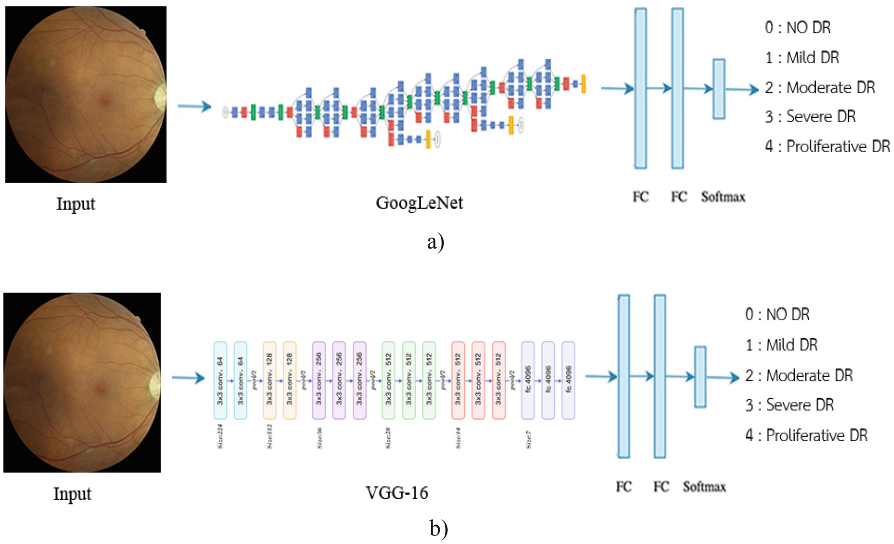

Fig. 9. Structure overview of the transfer learning method from moldel GoogLeNet (a). VGG-16 (b).

According to Fig. 9, Dense layers (FC) are shown in 2 layers, size 400 and 200 from the specified parameters where the cross-entropy loss is different. When comparing the Transfer Learning model with other models, the lowest cross-entropy loss of the Transfer Learning GoogLeNet model using Sigmoid as the Activation Function was 1.083 while the Transfer Learning VGG-16 model using ReLu as the Activation Function was 0.943. After adjusting the Learning Rate, it was found that the Transfer Learning

GoogLeNet model using Sigmoid as Activation Function and Learning Rate = 0.01 gave the lowest cross-entropy loss at 1.083 and Transfer Learning VGG-16 model using ReLu as Activation Function and Learning. Momentum Rate (Base LR = 0.03) provides the lowest cross-entropy loss at 0.843, which can be seen that the VGG-16 Transfer Learning Model yields lower cross-entropy loss than the GoogLeNet Transfer Learning Model. Thus, optimizing the Transfer Learning model is suggested. VGG-16 learning model, with some fine-tuning, the model can adjust the weight to better suit the data by fine-tuning at 3 $\times$ 3 convolution position onwards (Fig. 10). Using the Activation Function parameter (ReLu), the learning rate is momentum (Base LR = 0.03) and epoch equals to 300, which changes the dropout to Bath Normalize and uses L2 regularization in the dense layer.

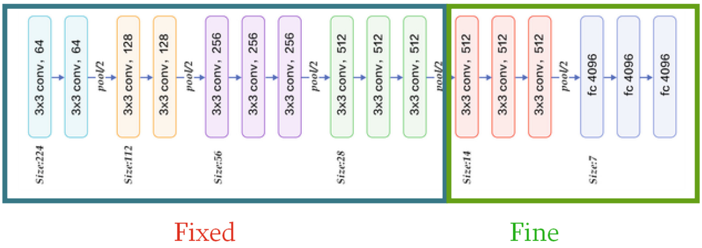

Fig. 10. Fine-tune of VGG-16

4 Results

The results of the experiment, when using the GoogLeNet model with Sigmoid Function as an Activation Function, the Learning Rate was 0.01, the percentage of accuracy and the Quadratic Weighted Kappa Score were 61.58% and 0.3731, respectively. As for the VGGNet-16 model, with fine-tuning, the smallest cross-entropy loss was 0.5761. The accuracy and the Quadratic Weighted Kappa Score were 71.65% and 0.5404, respectively. The proposed values were accurate for the classification of disease severity through GoogLeNet with VGG-16 models (Table 2).

Table 2. Measurement using confusion matrix in the classification of severity levels GoogLeNet and VGG-16.

Model	No DR	Mild DR	Moderate DR	Severe DR	Proliferative DR
GoogLeNet	85.86	19.17	44.31	45.07	50.06
VGG-16	79.89	42.38	48.00	58.13	60.71

Table 3. Comparison of the severity classification results of various model.

Method	Model	% Accuracy	Kappa score
Transfer learning	GoogLeNet	61.58	0.3731
	VGG-16	71.65	0.5404

When comparing the experimental results of the same data set, the VGG-16 Transfer Learning model provides a higher level of accuracy (71.65%) than GoogLeNet Model (61.58%) and when examining the Quadratic Weighted Kappa Score, the VGG-16 Transfer Learning model achieves a higher value of 0.5404 when compared with 0.3731 in GoogLeNet Model (Table 3).

5 Discussion

In the diabetic retinopathy examination, the DR and NO DR accuracy can be identified up to 85.86% and 79.89% of GoogLeNet and VGG-16, respectively. It can be seen that if you want to detect only diabetic retinopathy or not, GoogLeNet will be able to use it better. To detect the level of Diabetic Retinopathy, the VGG-16 version is recommended to detect DR. For diagnosing diabetic patients, and automatic patient screening would allow a considerable reduction in ophthalmologist workload with an increased level of accuracy in diagnosing diabetic eye disease. This would also allow ophthalmologists to assist distant hospitals with a lack of osteopathic expertise to perform DR examinations for remote patients. Nevertheless, this would greatly reduce the cost of examinations for hospitals and prospective patients.

Future research trends can use new techniques in the model to increase accuracy, but the time it takes to work is taken into account. You could try experimenting with red or blue images as inputs or experimenting with VGG19, ResNet, and Inception V3 to create the model.

6 Conclusion

In this research, CNN in DR is conducted based on the dataset from EyePACS. The Learning GoogLeNet and VGG-16 applications are compared to achieve the optimal timing for further improvement of the DR examinations. Based on the experimental results, the VGG-16 model with fine-tuning achieves a higher percentage of accuracy than GoogLeNet among with a better Quadratic Weighted Kappa score.

References

1. Barsegian, A., et al.: Diabetic retinopathy: focus on minority populations. Int. J. Clin. Endocrinol. Metab. **3**(1), 034–045 (2017)
2. Doshi, D., et al.: Diabetic retinopathy detection using deep convolutional neural networks. In: 2016 International Conference on Computing, Analytics and Security Trends (CAST) (2016)

3. Yau, J.W., et al.: Global prevalence and major risk factors of diabetic retinopathy. Diabetes Care **35**(3), 556–564 (2012)
4. Avidor, D., et al.: Cost-effectiveness of diabetic retinopathy screening programs using telemedicine: a systematic review. Cost Eff. Resour. Alloc. **18**, 16 (2020)
5. Xie, Y., et al.: Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a national programme: an economic analysis modelling study. Lancet Digital Health **2**(5), e240–e249 (2020)
6. Chatterjee, S., et al.: Cost of diabetes and its complications in Thailand: a complete picture of economic burden. Health Soc. Care Community **19**(3), 289–298 (2011)
7. Hsieh, Y.-T., et al.: Application of deep learning image assessment software VeriSee™ for diabetic retinopathy screening. J. Formosan Med. Assoc. (2020)
8. Alyoubi, W.L., Shalash, W.M., Abulkhair, M.F.: Diabetic retinopathy detection through deep learning techniques: a review. Inf. Med. Unlocked **20**, 100377 (2020)
9. Bellemo, V., et al.: Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study. Lancet Digital Health **1**(1), e35–e44 (2019)
10. Sayres, R., et al.: Using a deep learning algorithm and integrated gradients explanation to assist grading for diabetic retinopathy. Ophthalmology **126**(4), 552–564 (2019)
11. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
12. Teramoto, A., et al.: Automated classification of benign and malignant cells from lung cytological images using deep convolutional neural network. Inf. Med. Unlocked **16** (2019)
13. Nojiri, N., et al.: Discussion on machine learning and deep learning based makeup considered eye status recognition for driver drowsiness. Procedia Comput. Sci. **147**, 264–270 (2019)
14. LeCun, Y., et al.: Backpropagation applied to handwritten zip code recognition. Neural Comput. **1**(4), 541–551 (1989)
15. LeCun, Y., et al.: Handwritten digit recognition with a back-propagation network. In: Advances in Neural Information Processing Systems (1990)
16. Cohen, J.: Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. Psychol. Bull. **70**(4), 213 (1968)