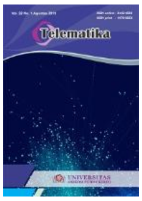

Available online at :
http://ejournal.amikompurwokerto.ac.id/index.php/telematika/

Telematika

Accreditated SINTA "2" Kemenristek/BRIN, No. 85/M/KPT/2020

UNIVERSITAS
AMIKOM PURWOKERTO

Convolutional Neural Network and Support Vector Machine Combination for Multi-Class Classification of Diabetic Retinopathy Based on Fundus Imaging

Tinuk Agustin[1], Moch. Hari Purwidiantoro[2], Ema Utami[3], Hanif Al Fatta[4]

[1,2] Informatics, STMIK Amikom Surakarta
[3,4]Magister of Informatics, Universitas Amikom yogyakarta
E-mail: tinukamikom@gmail.com[1], hariamikom@gmail.com[2], ema.u@amikom.ac.id[3], hanif.a@amikom.ac.id[4]

ARTICLE INFO

History of the article:
Received April 6, 2022
Revised July 21, 2022
Accepted August 15, 2022

Keywords:
Deep learning,
Machine learning,
Data split,
Confusion Matrix,
Computer-Aided Diagnosis

Correspondece:
E-mail:tinukamikom@gmail.com

ABSTRACT

Diabetic Retinopathy (DR) is a cause of blindness. Early detection has the potential to save the patient's vision. Reading fundoscopic photos requires more expertise and effort by the ophthalmologist. There are many visual similarities in lesions and only minor differences in the spatial domain. A computer-assisted automatic detection system is needed to assist medical experts in DR diagnosis and can reduce costs. This study proposes a combination method of Convolutional Neural Network (CNN) and Support Vector Machine (SVM) for the automatic classification of Diabetic Retinopathy (DR). The pre-train architecture Inception-V3 uses for feature extraction of input data. After training and getting the best model, the next is classification with SVM. Data augmentation techniques use to multiply and add variations to the dataset. Before the feature extraction stage, the dataset will process by separating the green channel from the RGB image. Next, the CLAHE will require increasing the contrast of the picture. This study aims to improve the performance in multi-class DR classification. The proposed model uses four classes of unbalanced and small datasets from retinal fundus images. This paper also compares the combined performance of CNN SVM with CNN Softmax based on the accuracy value to validate the results. Our experiments show that the combination of CNN SVM can increase the accuracy of auto-detection of DR severity up to 11.48% better than CNN softmax. The results showed that the pre-trained architectural model from the combination of Inception-V3 with SVM classification improves the accuracy extensively, even on small and unbalanced datasets.

INTRODUCTION

The initial stages of diabetic eye disease are Non-Proliferative Diabetic Retinopathy (NPDR). NPDR divide into mild, moderate, and severe NPDR. Ordinarily, in more severe stages, Proliferative Diabetic Retinopathy (PDR) causes retinal detachment (Brown et al., 2021). Clinically, the initial diagnosis of Diabetic Retinopathy (DR) is based on fundoscopy results. Fundus photographs use to observe retinal lesions at high resolution and to diagnose DR levels (Kwan, 2019). Retinal lesions that appear as early symptoms of DR are the presence of Microaneurysms (MA); Hemorrhage (HE); Hard Exudate (EX); Soft Exudate (SE) (Wysocka-Mincewicz et al., 2021). Such conditions make early detection of fundus images requires more expertise and effort by professional ophthalmologists, especially in remote areas with a limited number of ophthalmologists (Tsiknakis et al., 2021). Even though earlier detection of the NPDR phase is very significant to hold the worsening of the patient's visual function (Safi et al., 2018). The difference between NPDR and PDR is based on the presence of neovascularization, i.e. the growth of new retinal blood vessels due to preexisting ischemia. At any stage of DR, Diabetic Macular Edema (DME) may occur (Wysocka-Mincewicz et al., 2021). The development of artificial intelligence (AI) based Computer-Aided Diagnosis (CAD) system is very effective in assisting DR screening. AI can reduce the workload of reading fundoscopy photos and get faster and more accurate results. It is cheaper than the manual method (Tsiknakis et al., 2021).

Convolutional Neural Networks (CNN) is a deep learning method inspired by human eyesight, that can learn appropriate features directly from its data. CNN has outstanding feature extraction abilities, specifically in image recognition (Zhang et al., 2019). CNNs are very demanding on large computing and data resources like other deep learning methods. Since the discovery of GPU technology by Andrew et al. in 2009, the problem of the computational load has been resolved. However, medical images are very hard to obtain due to various factors, such as annotations, data size, unbalanced data, and relevant legal issues (Tajudin et al., 2022). All these factors make it hard to train an accurate and robust model free from overfitting. Overfitting is a critical problem in deep learning that often occurs. It occurs when models can learn patterns very well but can't predict them well on new datasets. Several manners did reduce overfitting, as in the study (Agustin & Sunyoto, 2020). They use augmentation data, dropout, and L-regulation. The result shows that using proper augmentation data can increase accuracy. Other research applies Early Stopping to training when the graph initiates to show overfitting (Reguant et al., 2021)(Dai et al., 2021).

In addition to the previously mentioned techniques, some researchers utilize transfer learning. Transfer learning trained with large data sets has proven capabilities and can transfer knowledge from other previously learned tasks (Goodfellow et al., 2016). Using transfer learning in the medical imaging domain will ultimately be much faster and more accurate than training a model from scratch. (Shanthi & Sabeenian, 2019) uses the modified Alexnet architecture to build the DR classification system into four severity levels. By processing 1190 fundus images and pre-processing using only the RGB green channel. He received an accuracy of 96.6%, specificity, precision, and sensitivity obtained 96.2%, 96.6%, and 95.6% respectively. (Hagos et al., 2020) adopted the Inception-V3 architecture to develop a DR classification system into five severity classes. This experiment divides an abnormal retina into four types. It used 1000 sets of images for normal retinas and 1000 images for abnormal retinas. The results of this experiment get an improvement in the accuracy value of 90.9% using softmax CNN.

Several experimenters combine CNN with different classification methods. In general, CNN stands employed for feature extraction and classification utilizing other classification methods. Such as classification using Multiple Kernel Learning (MKL) (Bold et al., 2019), Long Short-Term Memory (LSTM) (Liang et al., 2019), Xtreme Gradient Boosting (XGBoost) (Shete et al., 2020), and Support Vector Machine (SVM) (Taufiqurrahman et al., 2020)(Boral & Thorat, 2021). (Qomariah et al., 2019) conducted a study to examine the performance comparison of several popular transfer learning architectures. CNN feature extraction is combined with SVM classification to process 70 retinal fundus images and classify them into the retina normal and severe DR retina. They get good results in DR classification with 95.83% accuracy for inception-V3 architecture. Previous work was carried out by Sadiq, adopting the MobileNetV2 architecture to classify 5 DR severity levels. The system builds using the hybrid method of CNN and SVM to train highly unbalanced and limited data from the 2019 APTOS dataset. This hybrid system provides a solution to the data gap problem. The achievement they get promising performance with 85% accuracy

(Taufiqurrahman et al., 2020). Previous research has shown is not easy to get high performance in multi-class classification with limited and unbalanced datasets.

This study will discuss the classification severity of NPDR on a low and unbalanced retinal fundus image dataset. The combination method CNN SVM proposed for feature extraction and classification. SVM has the advantage process high-dimensional data without a significant decrease in performance and no need for dimensional reduction. Besides that, there is no overfitting term on it. The Inception-V3 architectural model was designed to work well on limited memory and computer resources. We propose a high-level image input understanding for CNN InceptionV3 architecture feature extraction. The features obtained will utilize as input for the classification on SVM. The raw data from the retinal fundus image carries only from the green channel, then increase the image using CLAHE. Resized needs as input for CNN extraction. Data augmentation uses only for more variation data. This preprocessing stage aims to get the informative medical image to easily further processing and automatic image evaluation stages. Medical images are hard to interpret, so the initial processing is consequential. Finally, linear SVM will use to replace the softmax function at the evaluation stage. This study will also compare the performance of the CNN Softmax method. The classification errors visualize and analyzed based on accuracy, sensitivity, and specificity using a multi-class confusion matrix.

RESEARCH METHODS

1. Dataset and Tools

The experiment for this study uses a google Colaboratory a free Jupyter Notebook environment that runs totally in the cloud with GPU access. The dataset takes from The Messidor database. The Messidor database (Decencière et al., 2014) was obtained from three ophthalmology departments and grouped or labeled by experts based on the DR's severity level. The images acquire in TIFF format with dimensions of 1440 * 960, 2240 * 1488, or 2304 * 1536 pixels. The total of images in this study is 984, with the details shown in Table 1.

Table 1. Data Sharing

No.	*Diabetic Severity*	*Total*
1	Normal	381
2	Mild	153
3	Moderate	196
4	Severe	254
	Total	984

2. Data Splitting

Data splitting uses to train and evaluate machine learning models and avoid overfitting. Data can split into training data, validation data, and test data. There are no specific guidelines for determining divided data. The exact ratio depends on several factors, including the size of the data and the number of predictors (Max Kuhn, 2019). But the most important is the data. The performance of the model increases directly with the amount of data. A random split algorithm and a reasonable choice of set ratio for training/validation and test data will achieve the same result (Birba, 2020). Usually, the highest ratio is used for the training data, followed by a balance between the validation and test sets, as in the experiment (Reguant et al., 2021). For this purpose, we divided the data into three subsets: training set, validation set, and test set with the ratio of 70:10:20. Data sharing was performed by a random split algorithm using the train_test_split provided by Sklearn from the Python Library, which offers various data processing features.

3. Proposed Research

The study adopted an experimental research method to analyze the CNN SVM when classifying the NPDR images. Several experiments to do to obtain high accuracy, sensitivity, precision, and specificity values. However, we also experimented with measuring the accuracy of the CNN softmax classification and proposed a comparison with the CNN SVM. This experiment aims to determine the best CNN feature extraction approach for the SVM's input classification. In this paper, the scenarios we use are as follows: The first scenario compares the accuracy of the CNN softmax classification results with the CNN SVM combination using transfer learning Inception V3. The second scenario compares the classification of CNN Softmax and the CNN SVM combination method on the pre-trained Inception V3 architecture. Figure 2 is an overview of the research flow. The first step is to enter the retinal fundus image into the preprocessing technique. Then, complete the training of the overall CNN model with a back-propagation algorithm. The softmax function will minimize the loss function in the back-propagation algorithm. Thus, trained CNN is utilized as a feature extraction to obtain better results. The best model will save for further processing. The features obtained from CNN's feature extraction model were formed into the data frame for input into SVM for NPDR classification.

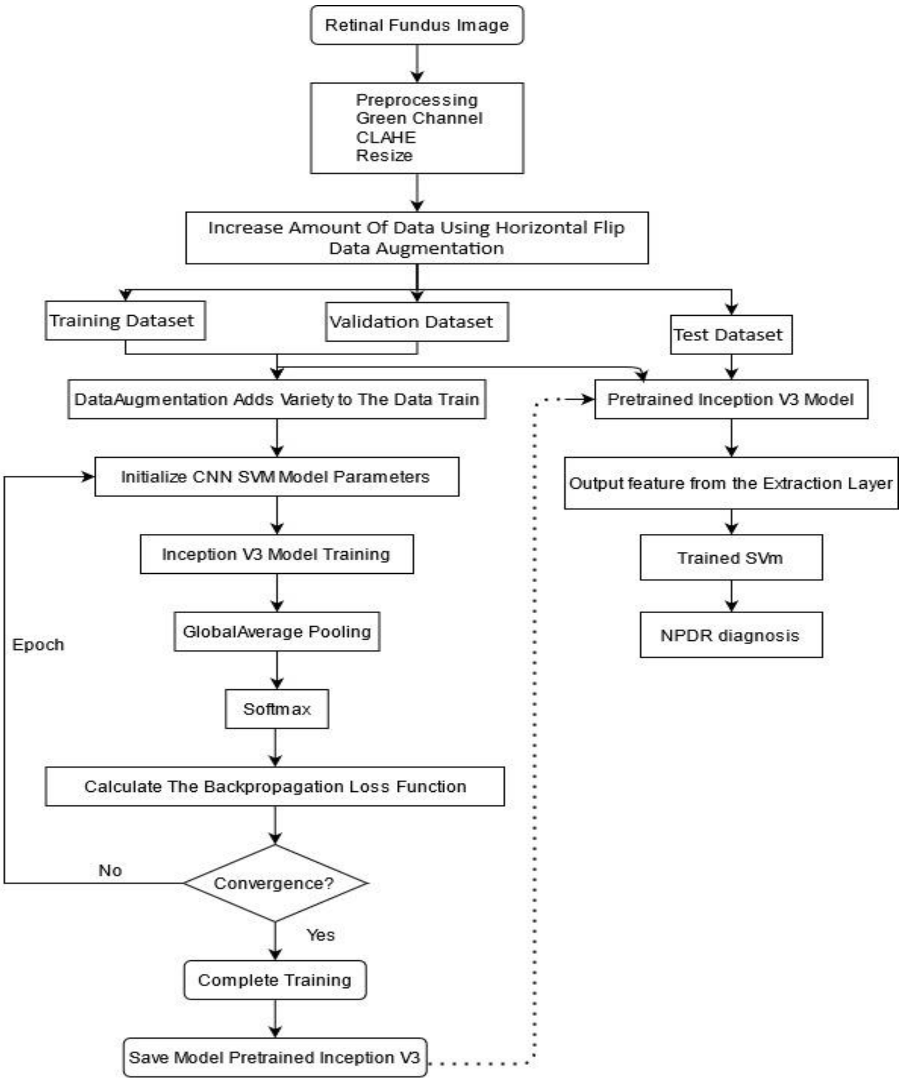

Figure 1. Research Flow

Selection of the proper process at the preprocessing stage can improve classification performance. Several steps conducted in the preprocessing process include:

a. Green channel extraction.

   The green channel chooses due to its ability to display maximum contrast between the background and the lesion. On the contrary, red and blue tend to contain more noise than green channels. In addition, the blood vessels can be seen more clearly in the green channel (Soomro et al., 2018).

b. CLAHE.

   The resulting green channel images will process using CLAHE to increase contrast. CLAHE can clarify the cross-section of blood vessels better than other contrast enhancement techniques (Erwin, 2020).

c. Resize.

   This step aims to equalize the image's pixel size to be processed so that suitable for the feature extraction process in the CNN architecture.

d. Data augmentation.

   This variation data enhancement approach includes a series of image manipulation operations such as random shift, flipping, random rotation, random zoom, and more. This process aims to artificially expand the data set to create new data (Wang & Perez, 2017). The selection of data augmentation techniques greatly influenced the experiment results (Agustin et al., 2020).

4. Build a Model

This study uses the InceptionV3 architecture model to design a system with memory limitations and holds computation costs. This system uses fewer parameters realized by 42 layers as a lower error rate is more efficient than VGGNet. This InceptionV3 model is an improved version of the previous version (Szegedy et al., 2015). Version 3 proposes the idea of convolutional factorization to reduce the number of parameters without reducing network efficiency. Figure 3 below shows the architecture of InceptionV3.

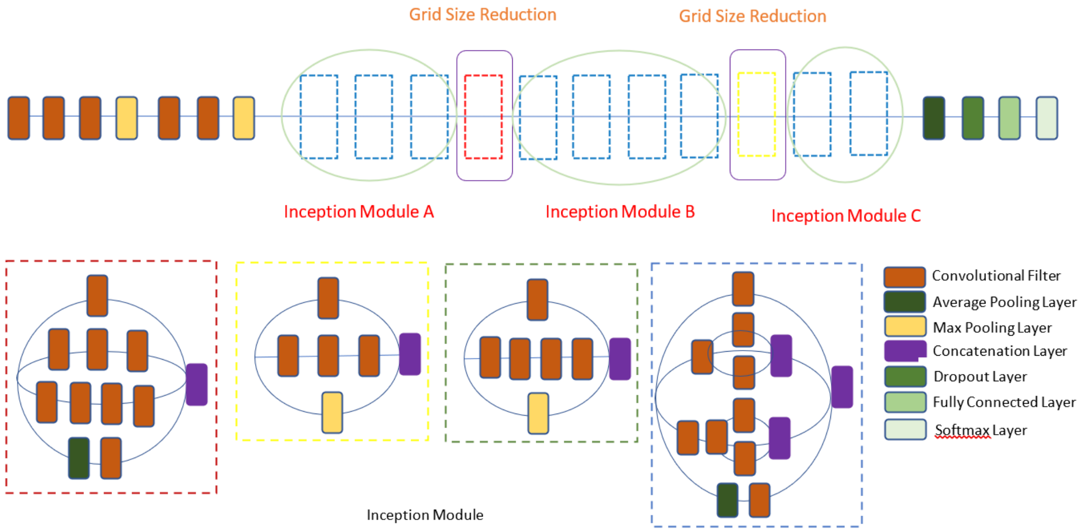

Figure 2. Architecture Inception V3 (Szegedy et al., 2015)

Transfer learning is then used to obtain feature vectors as input for the NPDR classification process into four levels of severity and compare which transfer learning results are best for classification. The process consists of removing the last layer, namely the fully-connected layer, and using the global average pooling (GAP) technology as a link for the classification layer, replacing the fully-connected layer. GAP effectively reduces the number of parameters so that the model's calculation time is also faster. Furthermore, the dropout is used as an adjustment to prevent overfitting. The linear SVM uses to replace softmax in the CNN classification to increase its accuracy and effectiveness. Linear kernels select since the data has a vast number of features. In general, the proposed model shows in Figure 4. The input layer is the input data with a size of 224*224px.

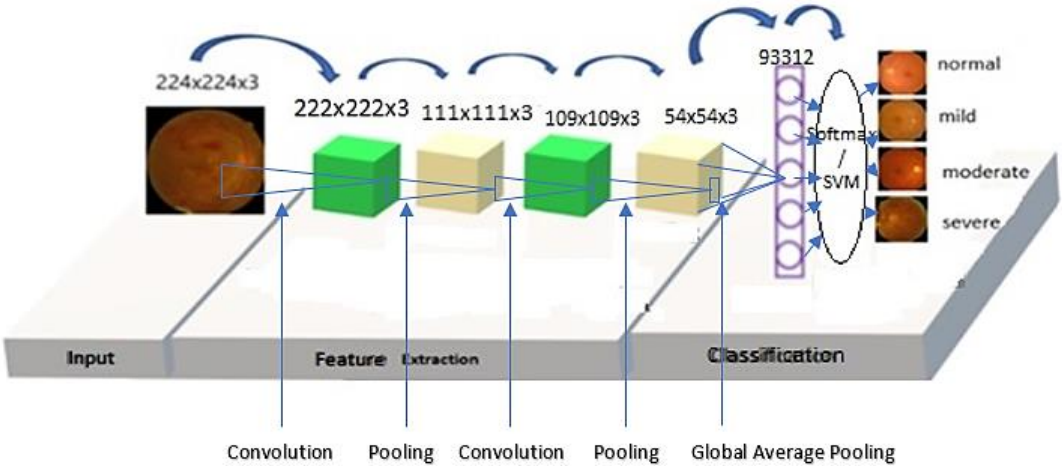

Figure 3. Built Classification Model Architecture

The convolution layer is applied as a filter to extract objects from the input image. The pooling layer aims to reduce the image size to arrange smoothly in the subsequent convolutional arrangement. Global Average Pooling strives to merge the final convolutional layer with a fully-connected layer. The objective layer is purely for classification. In our model, CNN could improve the network. Then, the developed model was saved and classified using SVM. The softmax function uses as the last layer of neural networks in multi-class classifications. SVM utilizes for the final classification engine.

5. Multi-class Confusion Matrix

The confusion matrix is a table used to measure the performance of the classification results from the model's predictions. This table presents the classification performance model based on actual or false test sets. True Negative (TN) means data correctly classified as negative or false outputs. True Positive (TP) shows the correctly classified data as positive or actual output. False Positive (FP) is data labeled No in the dataset but marked Yes in the prediction results. False Negative (FN) is data labeled Yes in the dataset but tagged No in the prediction results.

Table 2. Multi-Class Confusion Matrix

	Predicted Data				
	Classified Data	*Class 1*	*Class 2*	...	*Class n*
Actual Data	Class 1	X11	X12	…	X1n
	Class 2	X01	X11	…	X2n
	.	.	.	.	.
	.	.	.	.	.
	Class n	Xn1	Xn2	...	Xnn

Different than the 2-dimensional confusion matrix. In the multi-class confusion matrix, the values of TP, TN, FP, and FN determine based on Table 2. Equations (1), (2), (3), and (4) operate to calculate TP, TN, FP, and FN values (Manliguez, 2016).

$$TPall = \sum_{j=1}^{n} Xjj \quad (1)$$

$$FPi = \sum_{\substack{j=1 \\ j \neq i}}^{n} Xji \quad (3)$$

$$TNi = \sum_{\substack{j=1 \\ j \neq i}}^{n} Xj \sum_{\substack{k=1 \\ k \neq i}}^{n} Xk \quad (2)$$

$$FNi = \sum_{\substack{j=1 \\ j \neq i}}^{n} Xij \quad (4)$$

Explanation of the equation is TPall= the sum of all values in the diagonal confusion matrix, TNi = number of true negatives to i except row and column class *I*, *FPi* = number of false positives, *FNi* = number of false negatives *I*, *j* = index of row and column confusion matrix, *n* = total size of the row and column confusion matrix, *k* = matrix column index except class column index to *I*, *x* = value in matrix, *i* = column index of thr class matrix *I*.

The confusion matrix is the basis for calculating the accuracy, sensitivity, Precision, and specificity for each class and all classes can be seen in the equation below. The formula for calculating accuracy (ACC), sensitivity (SS), and specificity (SP) is in equations (5), (6), (7). Information is *N*= total number of testing entries, *i* = each class.

$$\mathrm{AC} = \frac{\mathrm{TP}all}{\mathrm{N}} \quad (5)$$

$$\mathrm{SSi} = \frac{\mathrm{TP}i}{\mathrm{TP}i+\mathrm{FN}i} * 100\%$$

$$\mathrm{SPi} = \frac{\mathrm{TN}i}{\mathrm{TNi}+\mathrm{FP}i} * 100\%$$

$$\mathrm{SS} = \frac{\mathrm{SS}all}{N} \quad (6)$$

$$\mathrm{SP} = \frac{\mathrm{SP}all}{\mathrm{N}} \quad (7)$$

RESULTS AND DISCUSSION

1. Experiment Configuration

The image preprocessing for each experiment includes several tasks: extracting the green channel, adjusting contrast with CLAHE, resizing the image size, augmenting data, and adding particular data variations to the data train. The specific training parameters for the test scenarios are in Table 3.

Table 3. Training Parameters using CNN

Parameter	*Information*
Optimizer	SGD
Loss	Categorical Crossentropy
Metric	Accuracy
Learning Rate	0,001
Epoch	50
Batch Size	56
Dropout	0.5

Epoch is the state when the dataset visits all neural network processes from start to finish and returns to the initial value for one cycle (back-propagation). The batch is a small unit of the epoch to be inserted into the computer. The learning rate is a measure of steps in each iteration carried out. While the SVM parameter used is linear SVM with a value of C = 50 and a value of penalty = l1.

2. Results The First Scenario Testing

The first experiment uses the CNN transfer learning model by utilizing a trained network of the InceptionV3 architecture. The only modification was replacing the fully-connected layer according to the data classification with four classes. The experiment results are in Table 4.

Table 4. Implementation of Transfer Learning Classification on Softmax and SVM

Inception V3	*Accuracy Obtained (%)*
Softmax	62,84

SVM	79,45

Table 4 shows the InceptionV3 transfer learning model has poor accuracy. We analyze that even though transfer learning from the ImageNet dataset improves the pattern recognition task's performance. But it's not work in the medical image dataset and does not significantly increase the pattern. This view is mainly due to the characteristics of the medical image dataset, which is very different from the previously trained data set. In contrast, using transfer learning architecture for the classification method by combining CNN SVM gives better results than CNN softmax. Although the CNN SVM combination increase accuracy significantly, it is still possible to improve.

3. **Results of The Second Scenario Testing**

The second scenario experiment aims to test the inceptionV3 model and retrain the whole data set. We present the results in Table 5.

Table 5. Pretrained Comparison Model Inception V3 Classification on Softmax and SVM

Inception V3	*Accuracy (%)*
Softmax	86,49
SVM	97,97

In Table 5, the accuracy value obtained from the pre-trained model InceptionV3 shows an improved result when using CNN-SVM. The analysis reveals that the network can learn well from the presented dataset so that the accuracy of CNN-SVM can exceed the accuracy obtained from the CNN classification. This improvement gets by optimizing the feature extraction of CNN that is input for the SVM classification process. The following steps in this scenario are the data extracted using a pre-trained InceptionV3 model architecture. Then, the results of the trained model are stored. The feature extraction results from CNN were reserved in the data frame and used as input for the SVM classification. Figure 5 presents the feature extraction results from the InceptionV3 model stored in the data-frame.

	feature0	feature1	feature2	feature3	feature4	...	feature1020	feature1021	feature1022	feature1023	label
0	0.383179	0.000000	0.647813	0.000000	0.616018	...	0.114113	0.208847	0.0	0.00000	3
1	0.000000	1.450444	1.541693	0.000000	0.140954	...	0.137513	0.060991	0.0	0.00000	3
2	0.582670	0.724713	0.000000	0.000000	1.981838	...	0.000000	0.213872	0.0	0.00000	0
3	0.000000	0.953972	0.667996	0.000000	0.000000	...	0.376218	0.000000	0.0	0.00000	3
4	0.254908	0.505501	0.000000	0.000000	1.835364	...	0.000000	0.239543	0.0	0.00000	0
...	...	...	...	...	...	...	...	...	...	...	...
1372	0.000000	1.084264	0.708647	0.000000	0.000000	...	0.088020	0.000000	0.0	0.00000	3
1373	1.083936	0.000000	0.793680	0.814126	0.170334	...	0.000000	1.014617	0.0	0.14868	2
1374	0.000000	0.110200	0.948195	0.000000	0.006557	...	0.000000	0.000000	0.0	0.00000	3
1375	0.000000	0.700365	0.437416	0.000000	0.029421	...	0.084835	0.000000	0.0	0.00000	3
1376	0.000000	0.782559	0.680391	0.000000	0.284066	...	0.000000	0.000000	0.0	0.00000	3

1377 rows × 1025 columns

Figure 4. Result of Fiture Extraction CNN

4. **Comparison Graph of CNN Softmax Classification Results and CNN SVM Combinations**

Based on the graph in Figure 6, the trend shows that the classification using SVM could increase the accuracy of results compared to the accuracy obtained by CNN. CNN softmax has significantly improved accuracy. It happens when the network re-trained. The combined CNN SVM can increase to 12.16% more than the CNN softmax classification. Moreover, we can ignore the overfitting when using SVM classification.

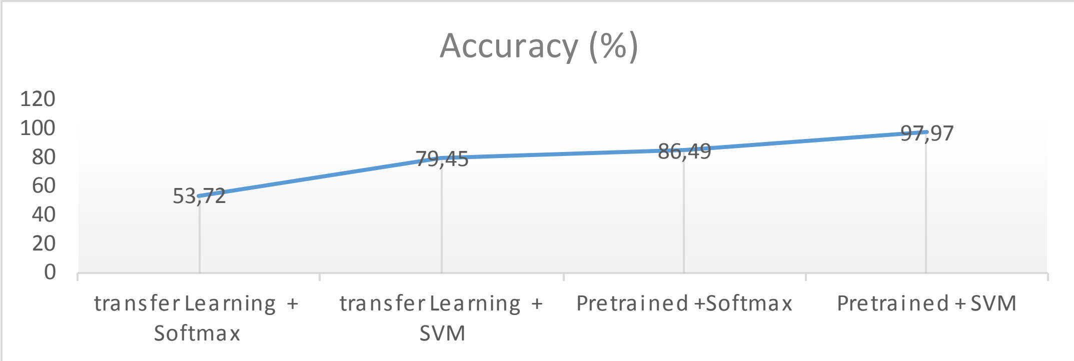

Figure 5. Comparison Classification of CNN Softmax Method and Combination Method of CNN SVM

5. Multi-Class Confusion Matrix Results

The confusion matrix calculation from the test results shows the best accuracy value is the combination method of CNN SVM with the InceptionV3 architectural model. Table 6 presents the results of the confusion matrix.

Table 6. Matrix Confusion

		Prediction Data				*Total*
	Level	*0*	*1*	*2*	*3*	
True	*0*	30	0	2	0	32
Disease	*1*	0	41	0	0	41
Level	*2*	1	1	72	1	75
Data	*3*	0	1	0	48	49
	Total	31	43	74	49	197

We used equations (1) to (7) to calculate each result value of the confusion matrix. Equations (1), (2), (3), dan (4) uses to calculate the TP, TN, FP, and FN values, respectively. The results obtained from these calculations use for calculating the ACC, SS, and SP values for each classification class based on equations (5), (6), and (7). The calculations' results in each category then find the average value. Table 7 presents the calculation results.

Table 7. Calculation of Matrix Confusion Results

	0	*1*	*2*	*3*	*Result (%)*
TP	30	41	78	48	-
TN	167	156	125	149	-
FP	1	2	2	1	-
FN	2	0	3	1	-
Accuracy					96,95
Sensitivity	93,75	100	96,25	97,79	96,95
Specificity	99,40	98,97	98,43	99,33	98,97

Based on table 7, the FN value has a higher result than FP. FN indicates a negative prediction error, where the patient is sorrowing from DR but is healthy. Therefore, the results need to interpretation with caution. In the real world, when the DR patient was predicted negatively and missed treatment, this condition could worsen the disease. Based on the result of equation (6), the sensitivity value shows a high of 96.93%. This result means that the system can separate each stage of severity very well, although still missed in some cases. In contrast to FN, FP implies prediction error in healthy patients suspected of having DR. For this reason, this paper does not include precision for the system feasibility assessment. It's not a risk for healthy patients diagnosed with DR because there will be further examination. Evaluation of this research is based on accuracy and sensitivity. The accuracy can measure whether the model performs can classify a normal retina from the NPDR retina very well. Sensitivity evaluates the test, where the patient's DR symptoms should not predict to be healthy. At the same time, the specificity indicates the ability of the test to ensure a healthy retina from an actual healthy retina. Based on the confusion matrix results in Table 7, the accuracy, sensitivity, and specificity results are 96.95%, 96.93%, and 98.97%.

6. Discussion

The following section discusses a comparison of the related to the NPDR classification using CNN transfer learning and the combination of CNN SVM with other similar studies. For the comparison, we referred to four recent studies and presented the results in Table 8.

Table 8. Comparison of Related Research

Author	*Model*	*Dataset*	*ACC(%)*	*SS(%)*	*Class*
(Qomariah et al., 2019)	Inception V3 + SVM	70	95,8	-	2
(Shanthi & Sabeenian, 2019)	Alexnet	1190	96,6	95,6	4
(Hagos & Kant, 2019)	Inception V3	2000	90,9	-	5
(Taufiqurrahman et al., 2020)	MobileNetV2+SVM	1805	85	-	5
Proposed method	Inception V3 + SVM	984	96,95	96,95	4

Our CNN SVM combination architecture model outperforms other models for similar classifications. The data set we used is smaller but well for multiclass DR classification. Augmentation techniques use only add variations to the training without adding data. As shown in table 8 for the DR 4 class classification, the accuracy we reach is almost the same as the study (Shanthi & Sabeenian, 2019) but we managed to increase the sensitivity value.

In our study, the CNN-SVM combination shows higher accuracy than CNN softmax even when involved in the multi-class NPDR classification. The results show higher accuracy using the pre-trained Inception V3 model for feature extraction and classification using SVM. These high results obtain as a method optimized for further use of feature extraction. Furthermore, retraining all transfer learning models to get high-accuracy results with good network performance also affects the results. The training model results were then saved and recalled for classification by replacing the fully-connected CNN layer with the SVM machine learning classification. The last step is removing the fully-connected layer. The features accepted from CNN feature extraction will be formed into a data frame for further processing to another machine learning classifier using SVM.

CONCLUSIONS AND RECOMMENDATIONS

This study proves that the combination method of CNN SVM could provide better accuracy in multi-class NPDR classifications than CNN Softmax. The high performance achieves by training the CNN network for its feature extraction using a pre-trained model from the Inception V3 architecture. After obtaining an excellent pre-trained CNN model, the model then stores for the following process. The last stage is eliminating the fully-connected layer (classification layer) and substituting it with the SVM classification method. The confusion matrix calculation results on the CNN model show the accuracy, sensitivity, and specificity values are 96.95%, 96.93%, and 98.97%. This study does not display the loss graph of CNN feature extraction results. Although the results are well and outperform other similar studies mentioned in this paper, the FN results obtained during the training are still too high. Supposedly in medical image detection, we should get a smaller FN value to provide a system with good performance.

REFERENCES

Agustin, T., & Sunyoto, A. (2020). Optimization Convolutional Neural Network for Classification Diabetic Retinopathy Severity. *2020 3rd International Conference on Information and Communications Technology (ICOIACT)*, 66–71. https://doi.org/10.1109/ICOIACT50329.2020.9332087

Agustin, T., Utami, E., & Fatta, H. Al. (2020). Implementation of Data Augmentation to Improve Performance CNN Method for Detecting Diabetic Retinopathy. *2020 3rd International Conference on Information and Communications Technology (ICOIACT)*, 83–88. https://doi.org/10.1109/ICOIACT50329.2020.9332019

Bold, N., Zhang, C., & Akashi, T. (2019). Bird Species Classification with Audio-Visual Data using CNN and Multiple Kernel Learning. *2019 International Conference on Cyberworlds (CW)*, 85–88. https://doi.org/10.1109/CW.2019.00022

Boral, Y. S., & Thorat, S. S. (2021). Classification of Diabetic Retinopathy based on Hybrid Neural Network. *Proceedings - 5th International Conference on Computing Methodologies and Communication, ICCMC 2021*, *Iccmc*, 1354–1358. https://doi.org/10.1109/ICCMC51019.2021.9418224

Brown, D. M., Wykoff, C. C., Boyer, D., Heier, J. S., Clark, W. L., Emanuelli, A., Higgins, P. M., Singer, M., Weinreich, D. M., Yancopoulos, G. D., Berliner, A. J., Chu, K., Reed, K., Cheng, Y., & Vitti, R. (2021). Evaluation of Intravitreal Aflibercept for the Treatment of Severe Nonproliferative Diabetic Retinopathy: Results from the PANORAMA Randomized Clinical Trial. *JAMA Ophthalmology*, *139*(9), 946–955. https://doi.org/10.1001/jamaophthalmol.2021.2809

Dai, L., Wu, L., Li, H., Cai, C., Wu, Q., Kong, H., Liu, R., Wang, X., Hou, X., Liu, Y., Long, X., Wen, Y., Lu, L., Shen, Y., Chen, Y., Shen, D., Yang, X., Zou, H., Sheng, B., & Jia, W. (2021). A deep learning system for detecting diabetic retinopathy across the disease spectrum. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-23458-5

Erwin. (2020). Improving Retinal Image Quality Using the Contrast Stretching, Histogram Equalization, and CLAHE Methods with Median Filters. *International Journal of Image, Graphics and Signal Processing*, *12*(2), 30–41. https://doi.org/10.5815/ijigsp.2020.02.04

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT press. http://www.deeplearningbook.org

Hagos, M. T., & Kant, S. (2019). *Transfer Learning based Detection of Diabetic Retinopathy from Small Dataset*. http://arxiv.org/abs/1905.07203

Hagos, M. T., Kant, S., & Bala, S. A. (2020). Automated Smartphone Based System for Diagnosis of Diabetic Retinopathy. *ArXiv*, 256–261.

Kwan, C. (2019). Object Tracking and Classification in Videos Using Compressive Measurements. *Proceedings of the 3rd International Conference on Vision, Image and Signal Processing*, 1–6. https://doi.org/10.1145/3387168.3387188

Liang, X., Qiao, X., & Li, Y. (2019). Obstructive Sleep Apnea Detection Using Combination of CNN and LSTM Techniques. *2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, 1733–1736. https://doi.org/10.1109/ITAIC.2019.8785833

Manliguez, C. (2016). *Generalized Confusion Matrix for Multiple Classes Generalized Confusion Matrix for Multiple Classes The total numbers of false negative (TFN), false positive (TFP), and true negative (TTN) for each class i will be calculated based on the Generalized*. *November*, 5–7. https://doi.org/10.13140/RG.2.2.31150.51523

Qomariah, D. U. N., Tjandrasa, H., & Fatichah, C. (2019). Classification of diabetic retinopathy and normal retinal images using CNN and SVM. *Proceedings of 2019 International Conference on Information and Communication Technology and Systems, ICTS 2019*, 152–157. https://doi.org/10.1109/ICTS.2019.8850940

Reguant, R., Brunak, S., & Saha, S. (2021). Understanding inherent image features in CNN-based assessment of diabetic retinopathy. *Scientific Reports*, *11*(1), 1–13. https://doi.org/10.1038/s41598-021-89225-0

Safi, H., Safi, S., Hafezi-Moghadam, A., & Ahmadieh, H. (2018). Early detection of diabetic retinopathy. *Survey of Ophthalmology*, *63*(5), 601–608. https://doi.org/10.1016/j.survophthal.2018.04.003

Shanthi, T., & Sabeenian, R. S. (2019). Modified Alexnet architecture for classification of diabetic retinopathy images. *Computers & Electrical Engineering*, *76*, 56–64. https://doi.org/10.1016/j.compeleceng.2019.03.004

Shete, M., Sabnis, S., Rai, S., & Birajdar, G. (2020). Diabetic Retinopathy Severity Detection using Convolutional Neural Network. *ITM Web of Conferences*, *32*(August), 01012. https://doi.org/10.1051/itmconf/20203201012

Soomro, T. A., Mahmood Khan, T., Khan, M. A. U., Gao, J., Paul, M., & Zheng, L. (2018). Impact of ICA-Based Image Enhancement Technique on Retinal Blood Vessels Segmentation. *IEEE Access*, *6*(c), 3524–3538. https://doi.org/10.1109/ACCESS.2018.2794463

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *07-12-June*, 1–9. https://doi.org/10.1109/CVPR.2015.7298594

Tajudin, N. M. A., Kipli, K., Mahmood, M. H., Lim, L. T., Awang Mat, D. A., Sapawi, R., Sahari, S. K., Lias, K., Jali, S. K., & Hoque, M. E. (2022). Deep learning in the grading of diabetic retinopathy: A review. *IET Computer Vision*, *May*. https://doi.org/10.1049/cvi2.12116

Taufiqurrahman, S., Handayani, A., Hermanto, B. R., & Mengko, T. L. E. R. (2020). Diabetic Retinopathy Classification Using A Hybrid and Efficient MobileNetV2-SVM Model. *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, *2020-Novem*, 235–240. https://doi.org/10.1109/TENCON50793.2020.9293739

Tsiknakis, N., Theodoropoulos, D., Manikis, G., Ktistakis, E., Boutsora, O., Berto, A., Scarpa, F., Scarpa, A., Fotiadis, D. I., & Marias, K. (2021). Deep learning for diabetic retinopathy detection and classification based on fundus images: A review. *Computers in Biology and Medicine*, *135*, 104599. https://doi.org/10.1016/j.compbiomed.2021.104599

Wang, J., & Perez, L. (2017). The Effectiveness of Data Augmentation in Image Classification using Deep Learning. *Computer Vision and Pattern Recognition*.

Wysocka-Mincewicz, M., Gołębiewska, J., Olechowski, A., & Szalecki, M. (2021). Diabetic retinopathy in children with type 1 diabetes—occurrence and screening using optical coherence tomography. *Life*, *11*(6), 1–9. https://doi.org/10.3390/life11060590

Zhang, Y. D., Dong, Z., Chen, X., Jia, W., Du, S., Muhammad, K., & Wang, S. H. (2019). Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation. *Multimedia Tools and Applications*, *78*(3), 3613–3632. https://doi.org/10.1007/s11042-017-5243-3