EMITTER International Journal of Engineering Technology
Vol. 10, No. 1, June 2022, pp. 195~216
DOI : 10.24003/emitter.v10i1.694

Classification of Ischemic Stroke with Convolutional Neural Network (CNN) approach on b-1000 Diffusion-Weighted (DW) MRI

Andi Kurniawan Nugroho[1,5], Dinar Mutiara Kusumo Nugraheni[2], Terawan Agus Putranto[3], I Ketut Eddy Purnama[1,4], Mauridhi Hery Purnomo[1,4,6]

[1] Electrical Engineering Department, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[2]Department of Computer Science, Universitas Diponegoro , Semarang, Indonesia
[3]RSPAD Gatot Subroto Presidential Hospital, Jakarta, Indonesia
[4] Department of Computer Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[5] Electrical Engineering Department, Universitas Semarang, Semarang, Indonesia
[6]University Center of Excellence on Artificial Intelligence for Healthcare and Society (UCE AIHeS), Indonesia
*Corresponding Author: hery@ee.its.ac.id

Received March 2, 2022; Revised April 3, 2022; Accepted May 5, 2022

Abstract

When the blood flow to the arteries in brain is blocked, its known as Ischemic stroke or blockage stroke. Ischemic stroke can occur due to the formation of blood clots in other parts of the body. Plaque buildup in arteries, on the other hand, can cause blockages because if it ruptures, it can form blood clots. The b-1000 Diffusion Weighted (DW) Magnetic Resonance Imaging (MRI) image was used in a general examination to obtain an image of the part of the brain that had a stroke. In this study, classifications used several variations of layer convolution to obtain high accuracy and high computational consumption using b-1000 Diffusion Weighted (DW) MR in ischemic stroke types: acute, sub-acute and chronic. Ischemic stroke was classified using five variants of the Convolutional Neural Network (CNN) architectural design, i.e., CNN1–CNN5. The test results show that the CNN5 architectural design provides the best ischemic stroke classification compared to other architectural designs tested, with an accuracy of 99.861%, precision 99.862%, recall 99.861, and F1-score 99.861%.

Keywords: Ischemic Stroke, Classification, CNN, b-1000 Diffusion-Weighted (DW) MRI, Accuracy.

1. INTRODUCTION

A stroke is a medical emergency because brain cells can die within minutes. Prompt treatment can minimize brain damage and possible

complications. In 2013, there are more than 2 million stroke patients in Indonesia, or 12 out of every 1000 people. About 80% of strokes are ischemic stroke [1]. Several factors are believed to cause ischemic injury, such as energy depletion and cell death. Brain imaging is one of the most significant techniques for the assessment of patients with ischemic stroke through Computerized Tomography (CT) scan and Magnetic Resonance Imaging (MRI). CT scan is more common because it is faster and cheaper, whereas MR has a much higher sensitivity of ischemic lesions [2]. Expert identification is still a common approach for studying stroke lesions on MRI images, as it is used to measure the size, shape, and volume of stroke lesions.

The computational analysis in image often assist doctor for diagnosis and help reduces the subjectivity of diagnosis, and also provide higher accuracy for aggressive treatments [3].

For stroke detection, machine learning approaches have been widely used. [4] research which detected lesions on MRI images with a hybrid approach. Unsupervised learning algorithms are used to segment brain images to produce probabilistic maps. Probabilistic map as input for Support Vector Machine (SVM) classification. Moreover, [5] research which analyzing stroke analysis using the SVM kernel and [6] research that set up logical rules for stroke with used data mining.

Research for Stroke detection with employing Deep learning techniques are rarely used. [7] research that compared several machine learning techniques to detects stroke (including Random Forests and Convolutional Neural Networks (CNN)). [8] research for detects pathology in brain images employs CNN and [9] research for stroke in electroencephalogram recordings with Deep Belief Network (DBN).

This research paper contributions are: first, we provide a public dataset consisting of b-1000 DW-MRI images of human brains from people with ischemic stroke, as well as healthy people. Another contribution of this work is to model the CNN model selection problem which results in high accuracy values in classifying ischemic stroke images.

Detection using Diffusion-Weighted Imaging (DWI) b-1000 required a radiologist to determine the period of obstruction (ischemic). DWI allows volume estimation for acute or severe conditions such as chronic (phase greater than 4 hours). The assessment of lesion volume is critical in the treatment of both acute and chronic patients, as it determines the ratio of hypoperfusion to core infarction [10]. Furthermore, acute and chronic lesions can be detected anatomically by combining the Region of Interest (ROI) function for each slice to form a volume function [11]. However, identifying ischemic stroke requires more experience on the part of a radiologist.

Identification of lesions has historically been viewed as an abnormality detection problem when an image of a healthy brain is observed, and harmful effects are detected in terms of tissue appearance [12]. However, the shape of the brain varies from one patient to another, and lesions can also damage brain structures. Furthermore, the MRI consist with a certain noise

and deformation, which leads to inadequate detection and classification. Hence, machine learning method are proposed to investigate the features of the training data and achieve high segmentation accuracy. Detection and classification are required, in the early stages of identifying an ischemic stroke.

This study describes another alternative for classifying using the Convolutional Neural Network (CNN) method on b-1000 Diffusion Weighted (DW) MRI images. In summary, the following are the main contributions of this paper.

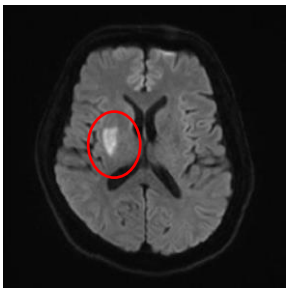
- Obtaining the b-1000 Diffusion Weighted (DW) MRI image dataset as a basis for input in the classification of ischemic stroke image levels in three classes, acute ischemic stroke, sub-acute ischemic stroke, and chronic ischemic stroke.
- Analyzing various CNN architectural designs to get the most appropriate CNN architecture in classifying ischemic stroke

2. RELATED WORKS

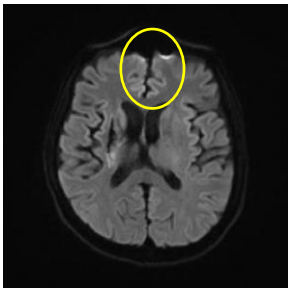
Several automatic and semi-automated detection studies were carried out to determine acute and chronic ischemic stroke lesions. Several methods are used to help solve clinical problems [13]. In general, researchers used data for ischemic stroke with a large area of blockage and only in the ventricles of the brain. The Random Forest method, for example, has been used in the literature to demonstrate good performance in brain tumor segmentation using handcrafted features [14][15][16][17]. The random forest method's quality is highly dependent on the characteristics extracted from an expert's prints. An algorithm is needed to determine the characteristics of the image to achieve excellent performance. For some tests, a semi-automated method of contouring at each stage can be used for manual analysis of each lesion. The procedure can be completed quickly if a large dataset is available. Quantitative volume determination is thus critical for ischemic lesions in order to save money or time [18][19]. Automatic and semi-automatic methods for completing visual assessments are introduced, and the dependence of experts on these evaluations is reduced. When the instruction is self-contained, this method is referred to as supervision.

To generate the trained model, most of the primitive feature recognition methods relevant to the context and their use are utilized. By utilizing the Convolutional Neural Network can eliminate this requirement. With the criteria, that because there is no feature extraction prerequisite. The learning network is used to complete automatic feature extraction and classification. From the above reasons, CNN is a method that is widely used for computer vision experiments. The success rate of using CNN is used for the graphics processing platform is very effective; optimized activation functions such as ReLU, and effective data augmentation techniques [20]. Biomedicine uses CNN a lot because of its end-to-end training and automated learning features. Chin et al. utilize CNN for ischemic stroke detection

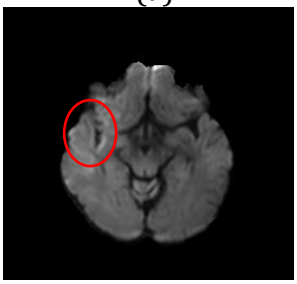
approach [21]. Another research employed Principal Component Analysis (PCA) combined with CNN for detecting brain tumors [22]. Diniz et al. introduced an approach to detect white matter lesions by combining CNN with SLIC0 clustering [23]. CNN was optimized using Particle Swarm Optimization to detect stroke lesions [24]. From prior research have trialed many methods for noticing brain lesions. However, it is rarely research which conduct an investigation for classifying the stroke using b-1000 Diffusion weighted imaging (b-1000 dwi). Therefore, this study conducts an experiment using multiple layer convolution classification method to determine the greatest accuracy in ischemic stroke types: acute, sub-acute and chronic using data b-1000 Diffusion weighted imaging (b-1000 DWI)



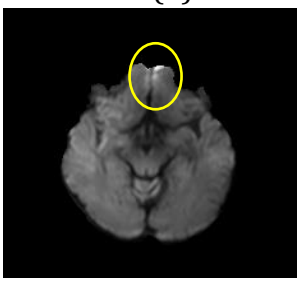
(a)



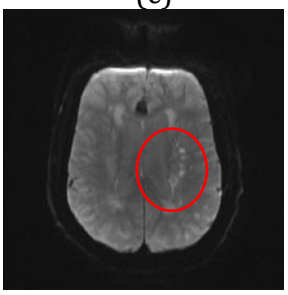
(b)



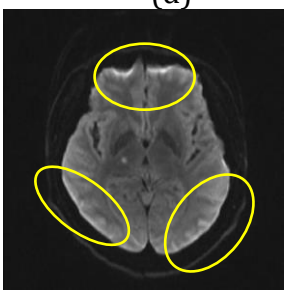
(c)



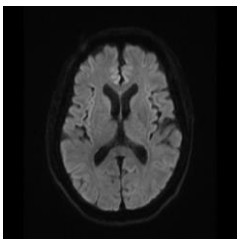
(d)



(e)



(f)



(g)

Figure 1. Example of a b-1000 DWI mode image for acute, ischemic, sub-acute, and chronic ischemia. Red circles represent acute, sub-acute and chronic ischemic lesions (a,c,e) and yellow circles represent artifacts for each class (b,d,f). For comparison, Figure 1(g) shows a normal b-1000 DW-MRI image.

Figure 1 shows the differences in b-1000 DW-MR images for acute, sub-acute, chronic and normal ischemic stroke classes. The red circle indicates the suspected ischemic stroke and the red circle indicates the image artifact.

3. ORIGINALITY

The problem of classification of acute, sub-acute and chronic ischemic lesions is very important in determining the early detection of stroke. The images are only distinguished by hypointense and hyperintense for each slice of the b-1000 MRI mode image, which is often required to detect this. Several researchers have used hyperintense notation in data sets while ignoring the distinctions between lesioned and normal tissue (Figures 1a, 1c, 1e) whereas artifact images are very difficult to distinguish from suspected lesions (Figure. 1b, 1d, 1f) [25].

This study describes another alternative for classifying using the Convolutional Neural Network (CNN) method on b-1000 Diffusion Weighted (DW) MRI images. In summary, the following are the main contributions of this paper.

- Obtaining the b-1000 Diffusion Weighted (DW) MRI image dataset as a basis for input in the classification of ischemic stroke image levels in three classes, acute ischemic stroke, sub-acute ischemic stroke, and chronic ischemic stroke.
- Analyzing various CNN architectural designs to get the most appropriate CNN architecture in classifying ischemic stroke

4. SYSTEM DESIGN

4.1 CNN design

The detection and recognition process often uses classifier methods such as the Hidden Markov Model [26], Support Vector Machine (SVM) [27], Artificial Neural Network [28], or the combined Adaptive Boost method [29]. Deep Learning (DL) methods have evolved in tandem with the development of hardware that supports big data processing and high-level computing. DL is a new scientific field within Machine Learning that focuses on the composition of nonlinear data transformations. Deep Neural Network (DNN) is one type of DL that has the advantage of having more than two layers. CNN is included in the type of DNN because it has a deep and multi-layered network. CNN is a supervised-feed forward DL that is a development of the Multilayer Perceptron (MLP) and is designed to process large-dimensional data, which is why it is widely used to process image/visual data [30]–[32].

CNN architectures such as VGG16 and VGG19 were introduced which have thirteen and sixteen convolution layers [33], ResNet50 and ResNet101 have 50 and 101 convolution layers [34], DenseNet 121, DenseNet 201, and DenseNet 169 have 121, 201, and 169 convolution layers [35], MobileNetV2 consists of 2 blocks where each block consists of 2 convolution layers [36], and MobileNetV3 in its original form consisting of 5 convolution layers has been developed to become simpler to become 2 convolutional layers [37]. The CNN architecture is divided into three layers based on their function, i.e., Convolutional Layer (CONV), Subsampling Layer (SUBS), and Fully Connected Layer (FC). The CNN architecture typically consists of multiple pairs of CONV and SUBS layers, followed by an FC layer.

The CONV layer is utilized to detect certain local features at all input image locations. The CONV layer also acts as a connecting layer, and converts the input data into a feature map that has been inflated with filters. The SUBS layer reduces the dimensions of the feature map by selecting the pixel values to output based on certain rules. The output of the CONV layer will be the input of the next CONV layer.

The ischemic stroke classification methodology was carried out using the CNN algorithm. The CNN used has the following main architecture: the CONV layer is equipped with a Rectified Linear Unit (ReLU) activation layer, SUBS layer using max pooling, and FC layer with Softmax activation function. The CNN method consists with first stage which is the backpropagation training stage, and the second stage is the feedforward image classification stage.

Table 1. Tuning Hyperparameter

Cost Function	Learning Rate (LR)	Optimizer	No. Epoch	Batch Size	Decay
categorical-cross-entropy	$1x10^{-4}$	Adam	30	32	LR/EPOCHS

The Table 1 shown in addition, part of the image pre-processing, hyperparameters represent an important part of the training. The best model for the classification of the ischemic stroke dataset was obtained after several configurations.

The system's initial input is a DWMRI b1000 mode image, which is a high-dimensional vector. A preprocessing step is usually required to condition the image and focus on the object to be classified. The preprocessing involves reimaging from 512x512 to 224x224, so that the .jpg data to be processed is focused on the ischemic stroke DWMRI infarct image.

In this CNN design, the activation function is designed using ReLU. The ReLU activation function is able to accelerate the convergence of the training process by increasing the sparsity of the network.

In the CONV layer, a convolution process is carried out on the input image using certain filters or kernels to produce a number of output features.

The CNN training process optimizes the filter matrix so that it can produce the most relevant features in the desired image classification process. The dimensions of these output features are large enough to make computation difficult. In this study, we use the MaxPooling algorithm on the SUBS layer, which is placed after the CONV layer, to reduce computational complexity and obtain a hierarchical set of image features. Max-pooling serves to improve translational invariance because it can reduce the size of the feature map by selecting the largest feature response. Then, normalization is carried out to simplify calculations, improve system robustness, and avoid errors in weighting initialization.

The final layer in the CNN is FC, which functions to classify based on the features obtained in the previous layer's calculation. If in the CONV layer ReLU is used as an activation function, then in the FC layer the Softmax function is used as an activation function. The Softmax function was chosen because it can reduce the possible error values generated by the cross-entropy objective function. In this study, the cross-entropy objective function is used because the problem to be solved is classification.

Table 2. Tested CNN Architecture Configuration

1st CNN	**2nd CNN**	**3thCNN**	**4thCNN**	**5thCNN**
CONV	CONV CONV	CONV	CONV CONV	CONV
ReLU	ReLU	ReLU	ReLU	ReLU
		Max Pooling		
			Max Pooling	Max Pooling
Max Pooling	Max Pooling	CONV		
			CONV CONV	CONV
		ReLU	ReLU	ReLU
		Max Pooling	Max Pooling	Max POOL
				CONV
				ReLU
				Max POOL
FC	FC	FC	FC	FC
Softmax	Softmax	SoftMax	SoftMax	SoftMax

The architectural design is presented in Table 2. CNN1 is the basic architecture of the CNN method. This study was conducted using five variations of the CNN architecture to obtain the most appropriate architecture in detecting the presence of ischemic stroke for acute, sub-acute,

and chronic stroke. Computing was carried out using a GPU GTX 1650 RAM 2 x 8 GB 2400 MHz DDR4.

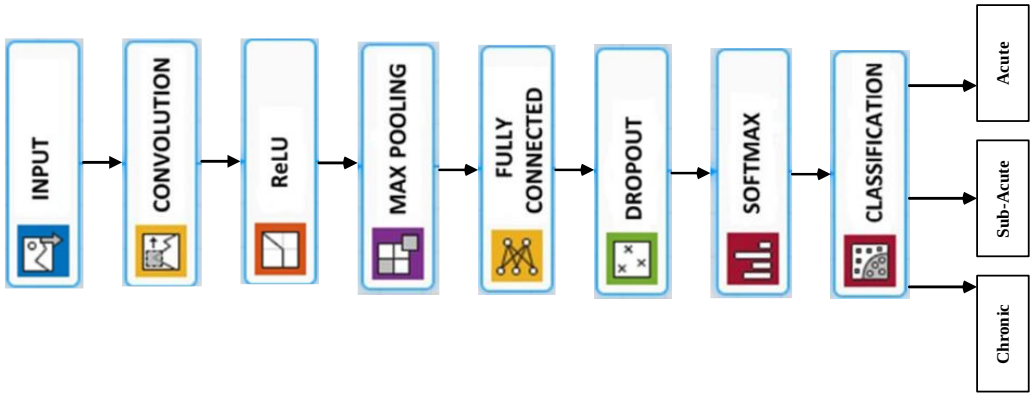


Figure 2. CNN1 model visualization

Figure 2 shows a visualization of the CNN1 architecture. This basic design can be developed into a CNN2 or CNN3 architecture. CNN2 can be expanded to CNN4, while CNN3 can be expanded to CNN5.

The CNN architecture is designed in this manner with the goal of determining the effect of the CONV layer and its activation function on classification accuracy. Comparisons can be made between these designs individually (CNN1, CNN2, CNN3, CNN4, CNN5) or between design groups to get the most appropriate type of design in classifying MR DWI image data for ischemic stroke.

4.2 Dataset

The dataset used is primary data taken from various hospitals. The image generated from the MRI machine is in the dicom file format measuring 512x512 for each slice in each patient. Conversion to png format is needed for the pre-processing process and converts to a size of 224x224 so that the computation of the program is not too heavy. The dataset for the trial process in this study used a collection of ischemic stroke images (acute, sub-acute, and chronic stroke) obtained from various hospitals in Indonesia. The dataset consisted of 21 patients for each class. Each patient consists of 16 slices for the axial plane. Addition of data is done by performing augmentations for each class such as rotation range, horizontal flip and width shift range.

Table 3. Specifications of the b-1000 MRI DWI Image Dataset

Process	Acute Image	Sub-Acute Image	Chronic Image
Training	768	768	768
Validation	192	192	192
Testing	240	240	240
Total	1200	1200	1200

The specifications for the number of datasets used in this study are shown in Table 3. Determination of the amount and specification of this data refers to the provisions of 80% training data/validation data and 20% test data.

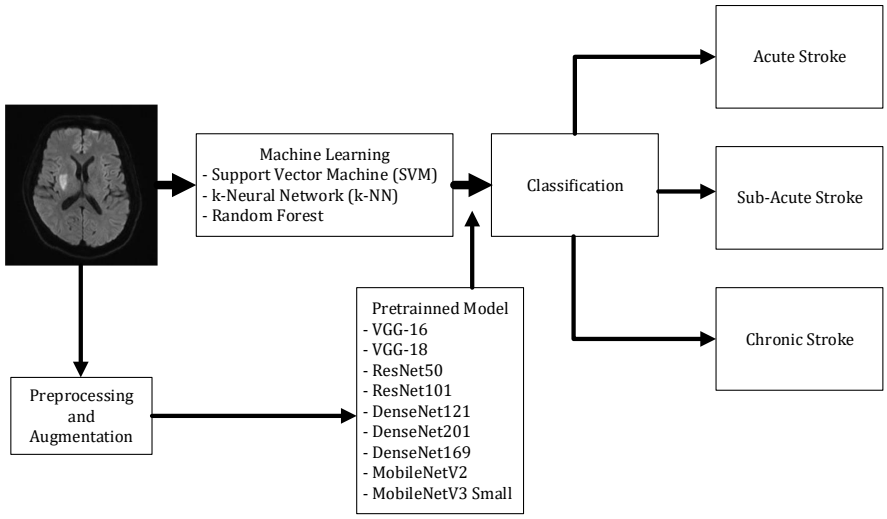


Figure 3. Diagram of research methodology

Figure 3 shows the research methodology for this study. First, the image processed in several stages pre-processing, data augmentation, and then training using a pre-trained algorithm: VGG-16, VGG-19, ResNet50, ResNet101, DenseNet121, DenseNet201, DenseNet169, MobileNetV2, MobileNetV3Small and Support Vector Machine (SVM), k–Neural Network (k-NN) and Random Forest classifications, and all algorithms tested on the data set test.

4.3 Pre-processing

One of the important steps in data pre-processing is resizing the b1000-DW MRI mode image as an input image for different algorithms. The b-1000 DWI MRI image is converted from dicom format to png format, with the original pixel size of 512x512 resized to and 224x224 pixels. All images were normalized according to the pre-trained model standard.

4.4 Data Augmentation

Actually, CNN performs better with large data sets. However, not all the data that in experiment can achieved a large data set. To address problem with large data set with the usage of CNN method, the data augmentation is the best solution. In addition, using data augmentation (rotation_range, horizontal_flip and width_shift_range) can improve the classification accuracy of deep learning algorithms. By adding existing data in deep learning models can improve the performance deep learning. Furthermore, few deep learning frameworks have data augmentation facilities embedded

into the algorithm. Therefore, for this paper, the experiment was used three augmentation strategies to generate new training sets (rotation, scaling, and translation).

4.5 Performance Matrix for Classification

Five CNN models, machine learning algorithms, and pre-trained model algorithms were trained and evaluated by comparing four performance metrics such as: accuracy, precision, recall, and F1-score [38]:

$$\text{Accuracy} = (TP+TN)/(TP+FN+FP+TN) \quad (1)$$

$$\text{Recall} = (TP) / (TP+FN) \quad (2)$$

$$\text{Precision} = (TP) / (TP+FP) \quad (3)$$

$$\text{F\_score} = 2x\ (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (4)$$

5. EXPERIMENT AND ANALYSIS

5.1 CNN Performance Analysis

Table 4. System Performance Analysis on Ischemic Stroke Classification

Method	Precision	Recall	F1-score	ACC (%)	Total Parameter	Computing Time (seconds)
CNN1	99.32	99.306	99.306	99.306	9.472.451	257.7
CNN2	99.725	99.722	99.722	99.722	18.947.043	266.1
CNN3	99.588	99.583	99.583	99.583	9.491.395	265.2
CNN4	99.055	99.028	99.028	99.028	20.547.619	265.8
CNN5	**99.862**	**99.861**	**99.861**	**99.861**	6.658.147	200.4

Table 4 shows the accuracy, precision, recall, and F1-score of each CNN1 - CNN5 architectural design in classifying b1000-DW MRI images from the 760 tested datasets using (1) - (5).

Based on the results in Table 4, it can be concluded that the CNN5 configuration showed the best results in classifying the b-1000 DWMRI image. The CNN5 model shows that in images with many features, a larger filtering process is needed. It is also seen that what is being tested is a complex dataset and the resulting increase in the percentage of accuracy compared to the simple CNN model.

5.2 Comparison with Other Classification Methods

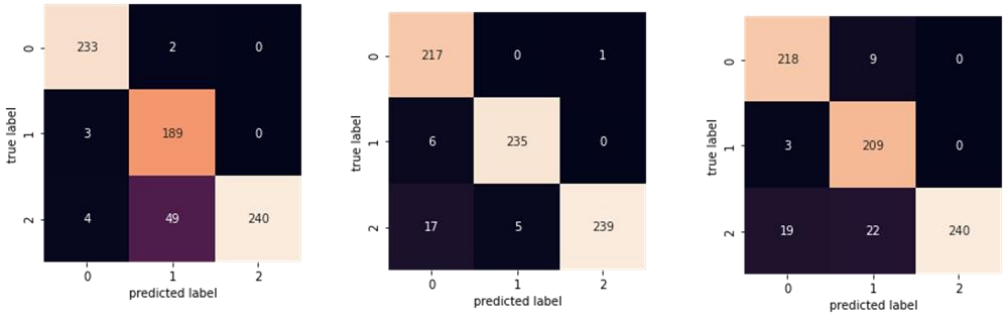
The following trials were conducted to compare the performance of the CNN5 design (as the best proven CNN design) with other classification methods, i.e., k–Neural Network (k–NN), SVM and random forest. For k–NN,

the D-MRI image dataset was tested using the k parameter of 3, while in the SVM method, the dataset was tested using a linear kernel. The use of random forest classifier with DW-MR data uses n_estimators=500. The test was conducted to classify the DW-MRI image dataset into acute, sub-acute, and chronic classes. The input data for the k–NN and SVM classifiers is the grayscale intensity value of each image pixel in the dataset.

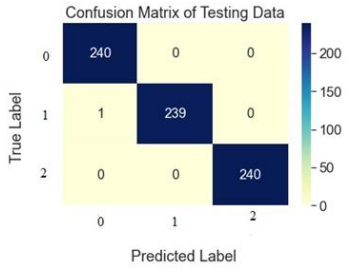
Table 5. Comparison of Performance of CNN – SVM – KNN – and Random Forest Methods on Ischemic Stroke Classification

Method	Precision (%)	Recall (%)	F-measure (%)	ACC (%)
Random Forest	92	93	92	92
SVM	93	93	93	93
Nearest Neighbors	96	96	96	96
CNN5	**99.862**	99.861	99.861	99.861

The test results are presented in Table 5. It was found that the four methods had good performance, with all of their accuracy, sensitivity, and specificity levels above 90%. The CNN5 method produces the highest performance compared to other classifier models. This shows that the CNN5 method can be implemented to classify DW-MRI images for ischemic stroke classification with good performance.



(a) Random Forest (b) Nearest Neighbors (k=3) (c) SVM



(d) CNN5

Figure 4. Comparison of the confusion matrix for the CNN 5 model with other classifications (Random Forest, Nearest Neighbors (k=3), and SVM

Figure 4 shown that based on the results of the tests, a confusion matrix for the classification of ischemic stroke can be created . The test results show that 233 typical acute image values were correctly predicted by the Random Forest confusion matrix classification, while 2 images were incorrectly predicted. Similarly, 189 sub-acute images and 240 chronic images were correctly predicted. Three sub-acute MRI images were not correctly predicted. Fifty-three chronic MRI images were not predicted correctly. The classification of the Nearest Neighbors (k=3) confusion matrix shows that 217 typical acute image values were predicted correctly, while 1 image was not correctly predicted.

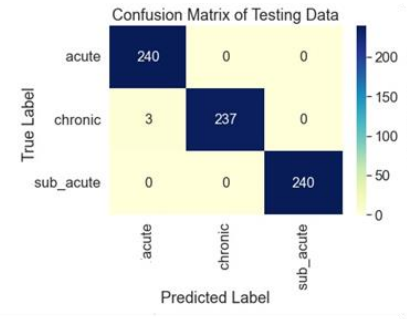
Similarly, 235 sub-acute images and 239 chronic images were correctly predicted. Six sub-acute MRI images were not correctly predicted. Twenty-two chronic MRI images were not correctly predicted. The classification of confusion matrix Support Vector Machine shows that 218 typical values of acute images were correctly predicted, while 9 images were not correctly predicted. Similarly, 209 sub-acute images and 240 chronic images were correctly predicted. Three sub-acute MRI images were not correctly predicted. Twenty-two chronic MRI images were not correctly predicted. The confusion matrix classifier of the proposed CNN5 model shows that 240 typical acute image values were correctly predicted. Similarly, 239 sub-acute images and 240 chronic images were correctly predicted. One sub-acute MRI image was not correctly predicted. The CNN5 model is the best in testing with testing data as many as 240 DWI images are able to predict almost all classes, only 1 image is actually sub-acute but predicted acute ischemic stroke

5.3 **Multi Grade Classification Performance**

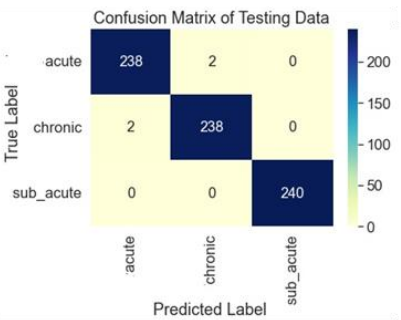
Table 6. Comparison of The CNN Method's Performance with the Transfer Learning Algorithm on Ischemic Stroke Classification

Method	Precision (%)	Recall (%)	F1_Score (%)	ACC (%)	Total Parameter	Computing Time (seconds)
VGG-16	99.588	99.583	99.583	99.583	17,602,883	333.72
VGG-19	99.444	99.444	99.444	99.444	24,747,587	1122.24
ResNet50	99.725	99.722	99.722	99.722	50,331,011	273
ResNet101	98.365	98.333	98.329	98.333	69,401,475	1395.15
DenseNet121	99.311	99.306	99.306	99.306	15,955,011	299.4
DenseNet201	99.723	99.722	99.722	99.722	34,579,523	303.3
DenseNet169	98.899	98.899	98.899	98.899	26,803,267	303.24
MobileNetV2	99.584	99.583	99.583	99.583	19,170,883	267.6
MobileNetV3 Small	99.584	99.583	99.583	99.583	19,170,883	267.3
CNN5	99.862	99.861	99.861	99.861	6,658,147	200.4

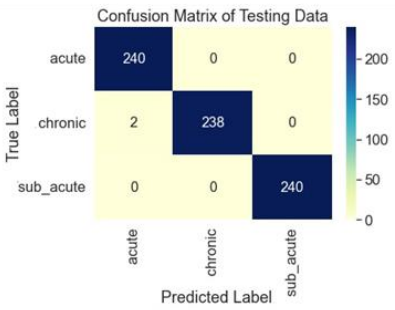
Table 6 shows a comparison of the performance of the nine pre-trained model. We trained nine pre-trained models on untrained data and validated them using a categorical classification, with grade 0 representing acute, grade 1 sub-acute, and grade 2 chronic. The results of the VGG-16 model were: precision 99.558%, recall 99.583%, F1-score 99.583%, accuracy 99.583%, and total parameters 17,602,883. VGG-19 showed promising results with 99.444% precision, 99.444% recall, 99.444% F1-score, 99.444% accuracy, and total parameters of 24,747,587. The results of the ResNet50 model were: precision 99.725%, recall 99.722%, F1-score 99.722%, accuracy 99.722%, and total parameters 50.331.011. The results of the ResNet101 model are: 98.365% precision, 98.333% recall, 98.329% F1-score, 98.333% accuracy, and 69.401.475 total parameters. The ResNet101 model has more parameters than the transfer learning model in this experiment. The results of the DenseNet121 model were: precision 99.311%, recall 99.306%, F1-score 99.306%, accuracy 99.306%, and total parameters 15.955.011. The results of the DenseNet201 model are: precision 99.723%, recall 99.722%, F1-score 99.722%, accuracy 99.722%, and total parameters 34,579,523. The results of the DenseNet169 model are 98.899% precision, 98.899% recall, 98.899% F1-score, 98.899% accuracy, and total parameters 26,803,267. The results of the MobileNetV2 and MobileNetV3Small models were: precision 99.584%, recall 99.583%, F1-score 99.583%, accuracy 99.583%, and total parameters 19.170.883. Finally, the CNN5 model produced 99.862% precision, 99.861% recall, 99.861% F1-score, 99.861% accuracy, and total parameters of 6,658,147.



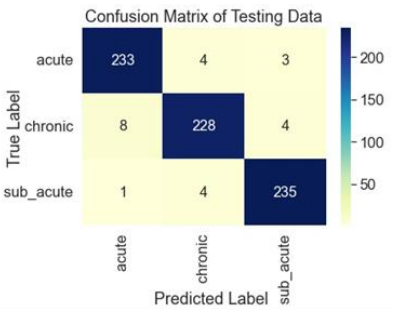
(a) VGG16



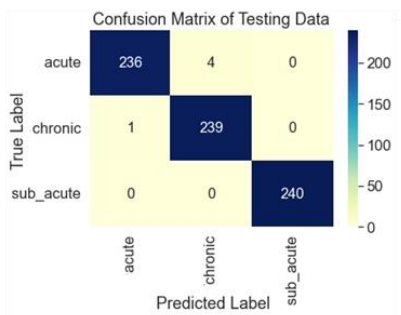
(b) VGG19



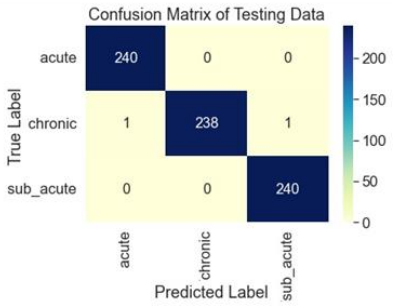
(c)ResNet50



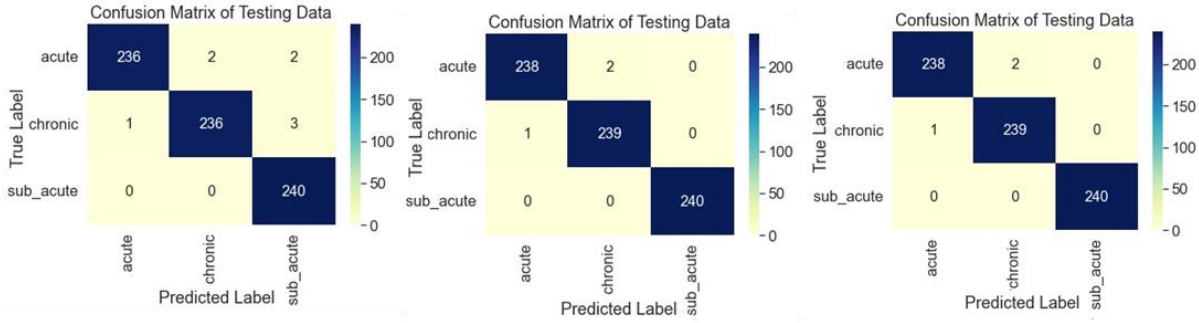
(d) ResNet101



(e) DenseNet121



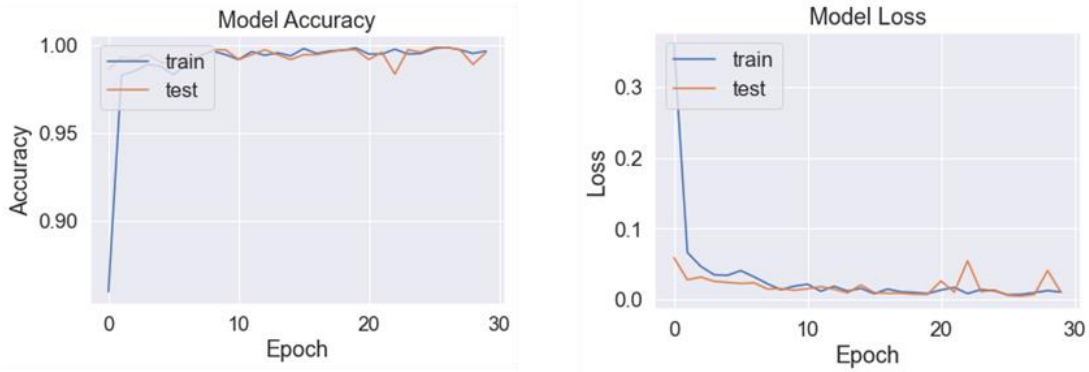
(f) DenseNet201



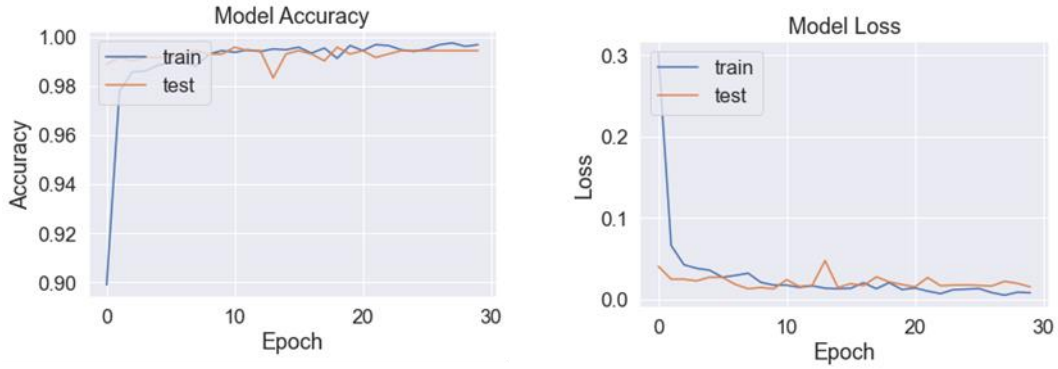
(g) DenseNet169 (h) MobileNetV2 (i) MobileNetV3Small

Figure 5. Confusion matrix for the classification of acute, sub-acute, and chronic ischemic stroke using nine pre-trained models

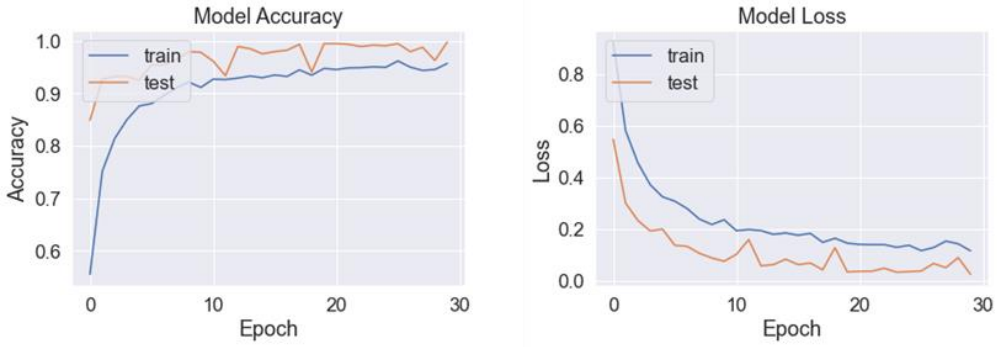
Figure 5 shows the confusion matrix that is used to determine the value of the TP, TN, FP, and FN labels. Then for the accuracy, precision, recall, and score of the F1 model, obtained by calculating the formula (1)–(4). Table 6 shows the performance metrics. From table 6 it is known that the CNN5 model has the best overall performance with a precision score of 99.862%, a recall score of 99.861%, an accuracy score of 99.861%, and an F1 score of 99.861%.



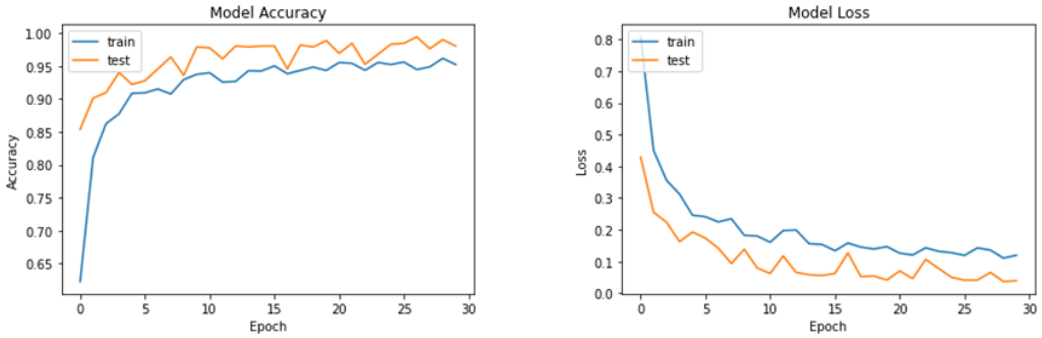
(a) VGG16



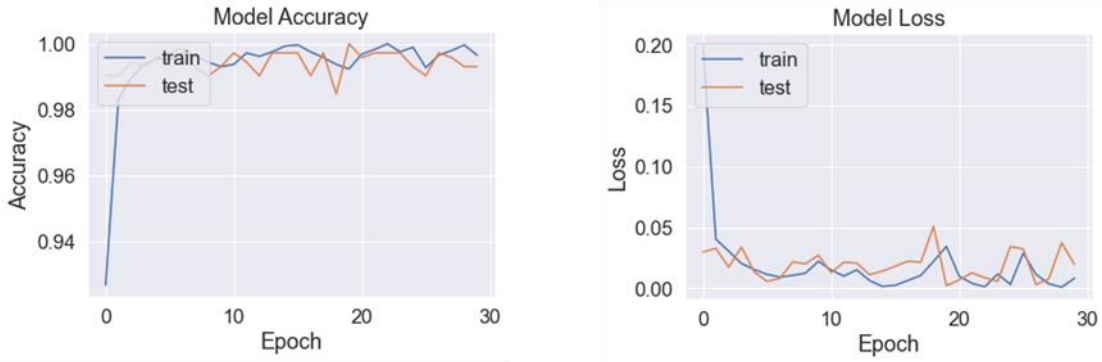
(b) VGG19



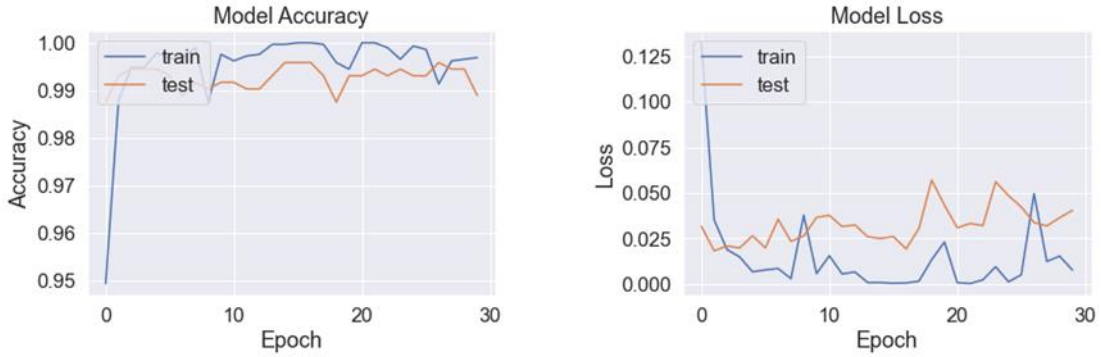
(c) ResNet50



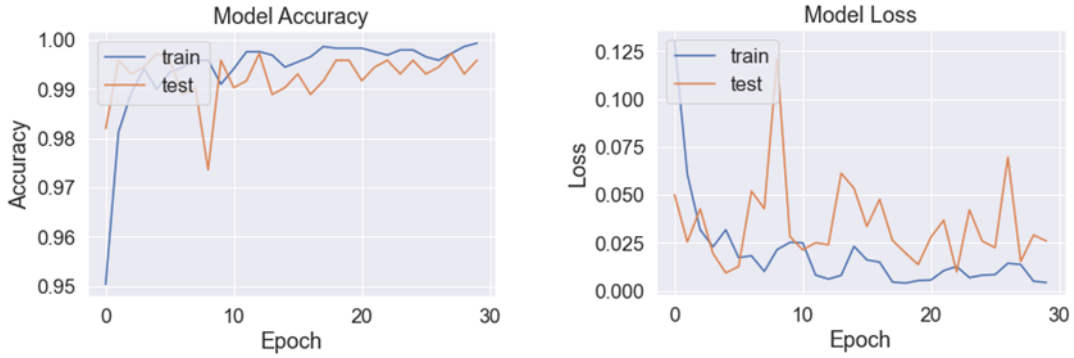
(d) ResNet101



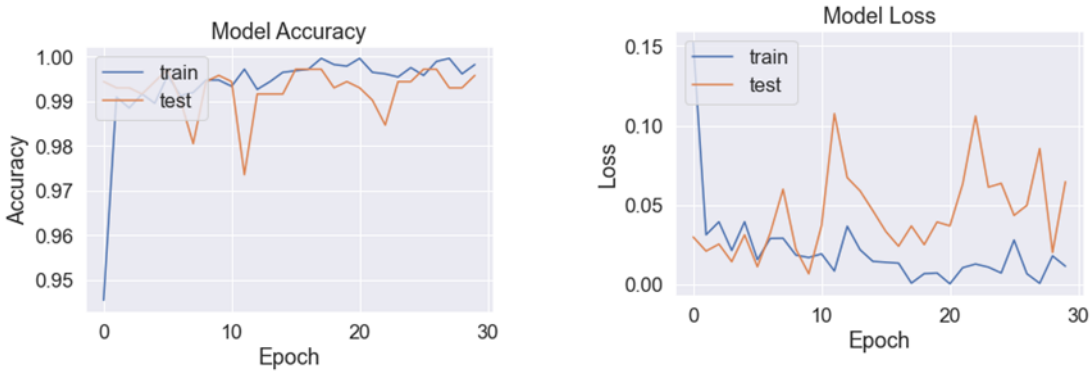
(e) DenseNet121



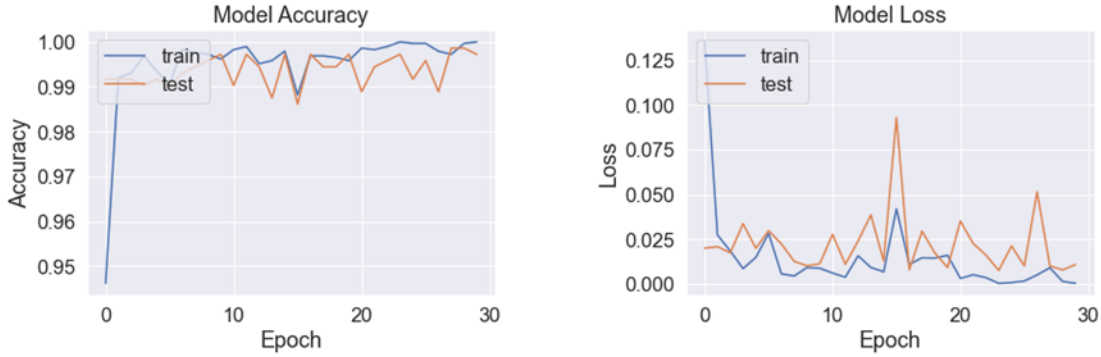
(f) DenseNet201



(g) DenseNet169



(h) MobileNetV2



(i) MobileNetV3Small

Figure 6. Accuracy and loss functions of all pre-trained models in training and validation data sets for multi-class classification

Figure 6 shows the results and model accurately predicting all random sample numbers as acute, sub-acute, or chronic. The VGG16 model shows the accuracy and plot loss function on the training data and validates the number of epochs. The instability in the VGG16 model arises from the significant difference between the training and validation data outputs. The results of the accuracy of the training data resulted in a value of 99.79%, and a loss of 0.0123 at epoch 26. The best validation data accuracy was 99.86%, and a loss of 0.0049 at epoch 28. These results indicate that the model has succeeded in classifying the data. validation. from learning information from training data and no overfitting. For the VGG19 model, the training accuracy is 99.69%, and the loss is 0.0100 at epoch 27. The validation accuracy is 99.58%, and the loss is 0.0169 at epoch 27.

These results indicate that the model was successful in classifying validation data from studying information from training data and there was no overfitting. For ResNet50 model, training accuracy was 100%, and loss was 0.0037 at epoch 30. Validation accuracy was 99.72%, and loss was 0.0084 at epoch 30. These results indicate that the model was successful in classifying validation data from studying information from training data and experiencing a little overfitting. For ResNet101 model, training accuracy was 96.18%, and loss was 0.1106 at epoch 29. validation accuracy was 99.44%, and loss was 0.0415 at epoch 27. These results indicate that the model was successful in classifying the validation data from studying information from the training data and experiencing a slight overfitting. For DenseNet121 model, training accuracy was 100%, and loss was 6.8103e-04 at epoch 26. The validation accuracy was 100%, and the loss was 0.0020 at epoch 27. These results indicate that the model had succeeded in classifying validation data from the information learned from the training data and there was no overfitting. For DenseNet201 model, training accuracy was 100%, and loss was 9.3803e-04 at epoch 16. Validation accuracy was 100%, and loss was 0.0022 at epoch 11. These results indicate that the model could successfully classify validation data from information learned from training data and slightly overfitting. For DenseNet169 model, training accuracy was 100%, and loss was 6.4445e-04 at epoch 18. Validation accuracy was 99.58%, and loss was 0.0205 at epoch 20. These results indicate that the model was successful in classifying validation data from information learned from training data and experiencing overfitting. For MobileNetV2_model model, training accuracy was 99.93%, and loss was 0.0012 at epoch 21. Validation accuracy was 99.86%, and loss was 0.0164 at epoch 24. These results indicate that the model was successful in classifying validation data from studying information from the training data and experiencing overfitting. For MobileNetV3Small model, training accuracy was 100%, and loss was 3.2804e-04 at epoch 29. Validation accuracy was 99.86%, and loss was

0.0072 at epoch 26. These results indicate that the model could successfully classify validation data from information learned from training data and experiencing overfitting.

Furthermore, the CNN 5 model has a layer structure that can classify the features of hyperintense and hypotensive ischemic stroke against the CNN 1-4 model. When compared to machine learning, it requires feature extraction and feature selection to produce high accuracy values. Meanwhile, with the pre-trained model, many features are produced due to the large number of convolution layers so that the system is Overfitting.

7. CONCLUSION

The study is contributed to Ischemic stroke classifications, using five variants of the Convolutional Neural Network (CNN) architectural design, i.e., CNN1–CNN5. The methodology research employed in this study is experimental ischemic stroke classification using five variations of the Convolutional Neural Network (CNN) architectural design, CNN1–CNN5. The experimental results show that the CNN5 architectural design provides the best ischemic stroke classification results, compared to other architectural designs tested. With the CNN5 architecture, the accuracy is 99.861%, precision is 99.862%, recall is 99.861, and F1-score is 99.861%. In addition, this study found that the results of comparisons with transfer learning algorithm methods such as VGG-16, VGG-19, ResNet50, ResNet101, DenseNet121, DenseNet201, DenseNet169, MobileNetV2, MobileNetV3Small and the classification of Support Vector Machine (SVM), k–Neural Network (k -NN) and Random Forest show that CNN5 has superior performance in terms of accuracy, precision, recall, and F-score. This study has not used hyperacute stroke data, so for further research, it suggested to conduct an experiment using hyperacute stroke data trialed with CNN5.

Acknowledgments

The author would like to thank the Ministry of Research, Technology, and Higher Education of the Republic of Indonesia for supporting this research through the Indonesian Education Scholarship (BPPDN). Furthermore, the authors also gratefully to the Gatot Subroto Army Hospital, the Department of Radiology, the Universitas Airlangga Hospital for providing MRI data for stroke patients, and the University Center of Excellence on Artificial Intelligence for Healthcare and Society (UCE AIHeS). In addition, this study was partially funded by the Education Fund Management Institute (LPDP) under the Innovative Productive Research Grant (RISPRO) scheme - Invitation 2019, contract number: PRJ-41/LPDP/2019.

REFERENCES

[1] Indah Permata Sari, **Faktor-Faktor yang Berhubungan dengan**

Terjadinya Stroke Berulang pada Penderita Pasca Stroke, Universitas Muhammadiyah Surakarta, 2015.

[2] A. K. Nugroho, T. A. Putranto, I. K. E. Purnama, and M. H. Purnomo, **Multi Segmentation Method for Hemorraghic Detection**, *2018 Int. Conf. Intell. Auton. Syst.*, pp. 62–66, 2018.

[3] E. R. da Silva, **Ambiente virtual colaborativo de diagn ´ ostico a dist ˆ ancia integrado a ferramentas de manipulac**¸ ˜ ao de imagens," Universidade Federal de Pernambuco, 2010.

[4] A. D. Guo, J. Fridriksson, P. Fillmore, C. Rorden, H. Yu, K. Zheng and S. Wang, **Automated Lesion Detection on MRI scans Using Combined Unsupervised and Supervised Methods**, *BMC Med. Imaging*, vol. 15, pp. 1–21, 2015.

[5] and A.-B. M. S. N. Farid, B. M. Elbagoury, M. Roushdy, **A Comparative Analysis for Support Vector Machines for Stroke Patients**, in *WSEAS Proceedings of the 7th European Computing Conference*, 2013, pp. 71–76.

[6] and P. J. T. Mroczek, J. W. Grzymała-Busse, Z. S. Hippe, **A Machine Learning Approach to Mining Brain Stroke Data**, *Springer Berlin Heidelb.*, pp. 147–158, 2012.

[7] C. S. O. Maier and and H. H. oder, N. D. Forkert, T. Martinetz, **Classifiers for Ischemic Stroke Lesion Segmentation : A Comparison Study**, *PLoS One*, vol. 10, pp. 1–16, 2015.

[8] and P. J. M. Havaei, N. Guizard, H. Larochelle, **Deep Learning Trends for Focal Brain Pathology Segmentation in MRI**, *arXiv*, vol. abs/1607.0, 2016.

[9] and N. A. G. Altan, Y. Kutlu, **Deep Belief Network Based brain Activity Classification Using EEG From slow Cortical Potentials in Stroke**, in *Proceedings of the International Conference on Advanced Technology & Sciences*, 2016, pp. 233–239.

[10] A. Wouters, P. Dupont, B. Norrving, and R. Laage, **Prediction of Stroke Onset Is Improved by Relative Fluid-Attenuated Inversion Recovery and Perfusion Imaging**, *Stroke*, pp. 2559–2564, 2016, doi: 10.1161/STROKEAHA.116.013903.

[11] M. P. P. ; S. T. ;Toan H. B. Visitsattapongse;Chuchart, **Automated Segmentation of Infarct Lesions in T1-Weighted MRI Scans Using Variational Mode Decomposition and**, *Sensor*, vol. 21, pp. 1–18, 2021, doi: https://doi.org/10.3390/s21061952.

[12] X. Liu, M. Niethammer, R. Kwitt, and M. Mccormick, **Low-Rank to the Rescue – Atlas-based Analyses in the Presence of Pathologies**, *HHS*, vol. 17, pp. 97–104, 2016, doi: 10.1007/978-3-319-10443-0_13.

[13] A. K. Nugroho, T. A. Putranto, M. H. Pumomo, and I. K. E. Purnama, **Semi Automatic Method for Basal Ganglia and White Matter Lesion Segmentation in MRI Images of Cronic Stroke Patients Using Adaptive Otsu**, *2018 Int. Conf. Comput. Eng. Netw. Intell. Multimedia, CENIM 2018 - Proceeding*, pp. 1–6, 2018, doi:

10.1109/CENIM.2018.8711285.

[14] Ellwaa A. et al, **Brain Tumor Segmantation Using Random Forest Trained on Iteratively Selected Patients**, 2016, doi: https://doi.org/10.1007/978-3-319-55524-9_13.

[15] L. Le Folgoc, A. V. Nori, S. Ancha, and A. Criminisi, **Lifted Auto-Context Forests for Brain Tumour Segmentation**, *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10154 LNCS, pp. 171–183, 2016, doi: 10.1007/978-3-319-55524-9_17.

[16] and L. S. L´aszl´o Lefkovits, Szid´onia Lefkovits, **Brain Tumor Segmentation with Optimized Random Forest**, in *MICCAI*, 2016, vol. 1, pp. 88–99, doi: 10.1007/978-3-319-55524-9.

[17] M.-C. L. Bi Song, Chen-Rui Chou, Xiaojing Chen, Albert Huang, **Anatomy-Guided Brain Tumor Segmentation and Classification**, in *International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, 2017, pp. 162–170, doi: https://doi.org/10.1007/978-3-319-55524-9_16.

[18] H. V. N. Z. Vemulapalli, **Cross-Domain Synthesis of Medical Images Using Efficient Location-Sensitive Deep Network**, in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015, pp. 1–8, doi: //doi.org/10.1007/978-3-319-24553-9_83.

[19] M. Ghafoorian *et al.*, **Location Sensitive Deep Convolutional Neural Networks for Segmentation of White Matter Hyperintensities,** *Sci. Rep.*, no. November 2016, pp. 1–12, 2017, doi: 10.1038/s41598-017-05300-5.

[20] M. Z. Abdelrahman Ellwaa, Ahmed Hussein, Essam AlNaggar, **Brain Tumor Segmantation Using Random Forest Trained on Iteratively Selected Patients**, 2016.

[21] A. C. Loic Le Folgoc, Aditya V. Nori, **Lifted Auto-Context Forests for Brain Tumour Segmentation**, 2016.

[22] A.-B. M. S. Heba Mohsena, El-Sayed A.El-Dahshan, El-Sayed M.El-Horbaty, **Classification Using Deep Learning Neural Networks for Brain Tumors**, *Futur. Comput. Informatics J.*, pp. 68–71, 2018.

[23] E. L. G. Pedro Henrique BandeiraDiniz, Thales Levi Azevedo Valente, João Otávio Bandeira Diniz, Aristófanes Corrêa Silva, Marcelo Gattassa , Nina Ventura, Bernardo Carvalho Muniz, **Detection of White Matter Lesion Regions in MRI Using SLIC0 and Convolutional Neural Network**, *Comput. Methods Programs Biomed.*, vol. 167, pp. 49–63, 2018.

[24] and V. H. C. D. A. D. R. Pereira, P. P. R. Filho, G. H. De Rosa, J. P. Papa, **Stroke Lesion Detection Using Convolutional Neural Networks**, 2018.

[25] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, and J. Winn, **The P ASCAL Visual Object Classes Challenge : A**

Retrospective, *Int. J. Comput. Vis.*, vol. 111, no. 1, pp. 98–136, 2015, doi: 10.1007/s11263-014-0733-5.

[26] P. Ambrosini, I. Smal, D. Ruijters, W. J. Niessen, A. Moelker, and T. Van Walsum, **A Hidden Markov Model for 3D Catheter Tip Tracking with 2D X-ray Catheterization Sequence and 3D Rotational Angiography**, *IEEE Trans. Med. Imaging*, vol. 0062, no. c, pp. 1–11, 2016, doi: 10.1109/TMI.2016.2625811.

[27] R. Rokhana, **Classification of Biomedical Data of Thermoacoustic Tomography to Detect Physiological Abnormalities in the Body Tissues**, in *2016 International Electronics Symposium (IES) Classification*, 2016, vol. 2, pp. 60–65.

[28] N. Tamami, P. S. Wardana, R. Rokhana, and H. Hermawan, **Neural Network Classification of Supraspinatus Muscle Electromyography Feature Signal**, in *2017 International Electronics Symposium on Engineering Technology and Applications (IES-ETA)*, 2017, pp. 223–228.

[29] Y. Yamasari, S. M. S. Nugroho, D. F. Suyatno, and M. H. Purnomo, **Meta-Algoritme Adaptive Boosting untuk Meningkatkan Kinerja Metode Klasifikasi pada Prestasi Belajar Mahasiswa**, *JNTETI*, vol. 6, no. 3, pp. 333–341, 2017, doi: http://dx.doi.org/10.22146/jnteti.v6i3.336.

[30] M. H. Purnomo, **Klasifikasi Nyeri pada Video Ekspresi Wajah Bayi Menggunakan DCNN Autoencoder dan LSTM**, *JNTETI*, vol. 7, no. 3, pp. 308–316, 2018, doi: http://dx.doi.org/10.22146/jnteti.v7i3.440.

[31] A. Nasuha, T. A. Sardjono, and M. H. Purnomo, **Pengenalan Viseme Dinamis Bahasa Indonesia Menggunakan Convolutional Neural Network**, *JNTETI*, vol. 7, no. 3, pp. 258–265, 2018, doi: http://dx.doi.org/10.22146/jnteti.v7i3.433.

[32] S. E. Limantoro, Y. Kristian, and D. D. Purwanto, **Pemanfaatan Deep Learning pada Video Dash Cam untuk Deteksi Pengendara Sepeda Motor**, *JNTETI*, vol. 7, no. 2, pp. 3–9, 2018, doi: http://dx.doi.org/10.22146/jnteti.v7i2.419.

[33] W. Setiawan and F. Damayanti, **Layers Modification of Convolutional Neural Network for Pneumonia Detection,** *J. Phys. Conf. Ser.*, vol. 1477, no. 5, 2020, doi: 10.1088/1742-6596/1477/5/052055.

[34] H. Wu, M. Xin, W. Fang, H. M. Hu, and Z. Hu, **Multi-Level Feature Network with Multi-Loss for Person Re-Identification**, *IEEE Access*, vol. 7, pp. 91052–91062, 2019, doi: 10.1109/ACCESS.2019.2927052.

[35] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, **Densely connected convolutional networks**, *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 2261–2269, 2017, doi: 10.1109/CVPR.2017.243.

[36] M. Sandler, M. Zhu, A. Zhmoginov, and C. V Mar, **MobileNetV2: Inverted Residuals and Linear Bottlenecks**, in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 4510–4520.

[37] A. D. J. Haicheng Wang, Vineeth Bhaskara, Alex Levinshtein, Stavros Tsogkas, **Efficient Super-Resolution Using MobileNetV3**, in *Computer Vision - ECCV 2020 Workshops - Glasgow, UK, August 23-28, 2020*, 2020, pp. 87–102, doi: https://doi.org/10.1007/978-3-030-67070-2_5.

[38] H. P. A. Tjahyaningtijas, A. K. Nugroho, C. V. Angkoso, I. K. E. Purnama, and M. H. Purnomo, **Automatic Segmentation on Glioblastoma Brain Tumor Magnetic Resonance Imaging Using Modified U-Net**, *Emit. Int. J. Eng. Technol.*, vol. 8, no. 1, pp. 161–177, 2020, doi: 10.24003/emitter.v8i1.505.