Received: February 16, 2022. Revised: March 11, 2022.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Utilizing the Hepta Convolutional Layer Neural Network (HCL-NN) Based on a Multi Optimizer for the Classification of Brain Stroke MRI

Andi Kurniawan Nugroho[1,4] 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's Email: ketut@te.its.ac.id; hery@ee.its.ac.id

Abstract: Stroke is a condition in which the blood supply to the brain is cut off. This occurs due to the rupture of a blood vessel. In addition, this can lead to fatal conditions for the patient where causes damages in several areas of the patients' brain. The examination by health workers is generally carried out to receive an image of the part of the patient's brain that has a stroke disorder. Commonly, stroke detection is done using magnetic resonance imaging (MRI). The purpose of this study is to use convolutional neural network (CNN) as a method for classifying MRI stroke images. The data used in this study is the DWI stroke MRI image dataset 5,226 images. The data consisted with 1,742 normal images, 1,742 intra cerebral hemorrhage (ICH) images, and 1,742 acute ischemic images, which are grouped into 3 parts, training, validation, and testing. The images produced by the MRI are labeled manually by the doctor to obtain certain types of strokes. This study employed the multi optimizer which is Root Mean Squared Propagation (RMSprop), adaptive moment estimation (Adam), and stochastic gradient descent (SGD). This study compared between the scenario original CNN model and the scenario hepta convolution layer neural networ*k* (HCL-NN) which proposed by this study. Both of the scenario of CNN tested using magnetic resonance –diffusion weighted (MR-DWI) images. Transfer learning in the CNN model (inside the CNN library) was tested with the author's data for performance testing. Then, the results of the performance trials that have been carried out earlier are compared with the results of the original CNN and HCL-NN. The experiment shows by using scenario HCL-NN display data training accuracy value 99.23 percent and data testing accuracy 97.40 percent. The both values are achieved with using RMSprop optimizer. Furthermore, this study also tested the values of precision, Recall and F1-scores. The scores for testing the HCL-NN scenario 97.43, 97.40 and 97.40 are higher compare to the other scenarios (original CNN) and other transfer learning (RESNET, VGG19 and mobilenet V2). As a summary of this study, the proposed HCL-NN performed better than other CNN scenarios for classifying strokes for MR-DWI images.

Keywords: Classification, CNN, HCL-NN, Accuracy, Multi optimizer.

1. Introduction

The brain is the center of the nervous system that regulates and coordinates most of the movement, behavior, and functions of the human body. Damage to the brain results in the cessation of some functions of the human body (paralysis). One of the diseases that can damage the brain is stroke. Stroke or cerebrovascular injury (CVA) is a condition that occurs when the blood supply to the brain is cut off due to blockage or rupture of a blood vessel, resulting in the death of some areas of the brain. This disease is a health condition that can be fatal if not treated quickly [1].

According to WHO (world health organization) in 2012, 51% of deaths due to stroke worldwide are caused by high blood pressure (hypertension).

According to 2013 basic health research data, the prevalence of stroke in Indonesia is 12.1 per 1,000 population. Stroke has become the main cause of death in Indonesia. Stroke rates were significantly higher in men (7.1%) compared to women (6.8%) [2].

The most important factor causing a stroke is age because stroke often occurs in the elderly [3]. A stroke can cause paralysis on one side of the body. Paralysis usually occurs on the opposite side of the lesion in the brain because of cross-representation regulation by the brain. In addition, stroke also causes visual disturbances, aphasia (difficulty speaking or understanding speech), decreased consciousness, and dementia (impaired intellectual function) [4].

One way to diagnose stroke is computed tomography (CT) Scan and MRI (magnetic resonance imaging). CT scan is performed using a series of X-rays to produce clear and detailed images of the patient's brain. CT scans can show the presence of bleeding in the brain, ischemic stroke, tumors, and various other health conditions. The doctor may also inject a dye into the bloodstream to get a better view of the blood vessels in the neck and brain. If a patient is thought to be having a stroke, a CT scan can show the doctor the type of stroke the patient is having [5].

CT scans are thought to be faster than MRIs, allowing patients to receive effective treatment as soon as possible. An MRI is usually done using radio waves and strong magnets to create clear and detailed images of the patient's brain. This method can detect brain tissue damaged by ischemic stroke and brain hemorrhage. Usually, this method is performed on patients who have various symptoms, so the location of the damage is still unknown. This method is also performed on patients who are recovering from a transient ischemic attack (TIA) or minor stroke. This test also provides more detailed images of brain tissue and allows for the detection of locations that would otherwise be invisible. A colored fluid may be injected into a vein to allow the doctor to see the arteries and veins and explain the flow of blood in the body [6].

Research to identify stroke has been carried out previously through CT image results using the statistical feature extraction feature to characterize the histogram with four parameters, i.e., mean, standard deviation, skewness, and kurtosis. Contrast limited adaptive histogram equalization (CLAHE) has also been used to clarify the image by providing a limit value for increasing image contrast. A study uses the statistical region merging method for image segmentation [7]. The next research is to detect and classify cerebral hemorrhage from CT scan images using the extreme learning machine artificial neural network method for classification. The image processing process uses gray scaling, scaling, and thresholding for image segmentation. Previous research done by Malathi et al. [7], classify strokes based on feature extraction for each type of stroke. The disadvantage of this approach is that the method cannot distinguish between types of acute ischemic stroke and Intra Cerebral Hemorrhage.

Another study has proposed the convolutional neural network (CNN) method. Convolutional neural network is a machine learning that has 1 training stage (supervised back-propagation). CNN is part of the neural network method, which works similarly to MLP (multi-layer perceptron). The convolutional neural network has previously been used to detect brain tumors through magnetic resonance imaging (MRI) images [8]. Subsequent research uses CNN to classify lung cancer through CT scans [9]. Subsequent studies use CNN for the classification of Alzheimer's disease through MRI images and MRI flares. From the study [8, 9], the researcher only identified only identify tumors and Alzheimer's, without identifying the presence of stroke in the CT scans images.

Therefore, this study proposes hepta convolutional network (HCL-NN) for classifying stroke that is based on convolutional neural network (CNN) method on MRI images. HCL-NN is a CNN model that has seven convolution layers to produce many extraction features on each layer on each MRI image in each class and classification accuracy value. The remainder of this paper is organized as follows: section 2 explains CNN architecture the proposed methodology for this study, and section 3 provides details of the experiment to test the process of the scenario models MR image datasets and the results. Meanwhile, this section discusses the result of the scenario model and comparison with the model of transfer learning architecture and section 4 summarises the result of this study.

2. Methodology

The proposed method for classifying stroke consists of several stages. These stages start from the collection of normal, ischemic stroke, and hemorrhagic stroke image data which will be used for training and test images. The labeling stage is used to divide the image into 3 classes. Scaling is used in the pre-processing stage to determine the pixel size used in brain image processing [10]. Then, the uniformity of the gray image is done using grayscale. The resize stage is used for the data input stage. To perform data training, data splitting is used to divide the data into three, i.e., training data, testing data, and validating

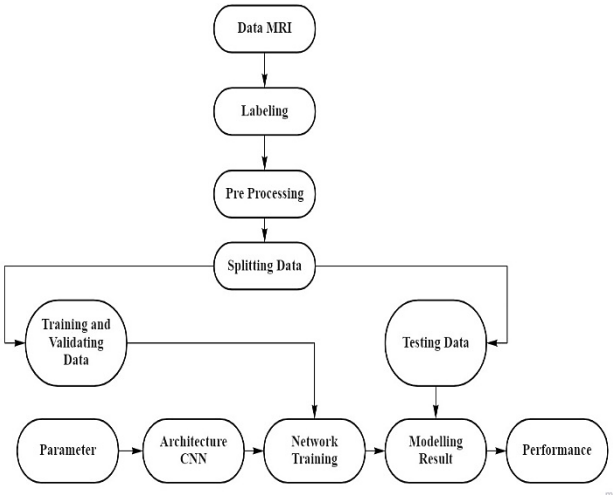

Figure. 1 MR stroke image classification process

data. The CNN process will continue by tuning the CNN parameters, such as the number of epochs, learning rate, so that the next step will be determined to calculate the performance value of the classification accuracy for each class.

2.1 Augmentation data

The image data used in this study is an MRI image obtained from some hospitals in Indonesia. The CT Scan image data obtained were 150 images, i.e., 50 normal images, 50 ischemic stroke images, and 50 hemorrhagic stroke images. The data will be used as training data and test data. To prevent overfitting, data is added by performing augmentations such as rescaling data, shearing, zooming, horizontal flip [11]. Data sharing is carried out with a ratio of 80 percent of training data (20% of training data is taken for validation) and 20 percent of testing data.

2.2 CNN Architecture

The stroke classification approach used in this study is the convolutional neural network (CNN) method with different layer modifications from previous studies. The CNN method uses 3 and 5 convolution layers with filter values of 32, 64, and 128. The classification process uses a flattening function, a fully connected layer, and a dense function. The CNN enhancement method modifies the number of convolutions in the CNN architecture. The trial process was carried out using an MRI-DWI image dataset on stroke patients for ischemic, hemorrhagic, and normal types obtained from various hospitals in Indonesia. The test scenario is carried out with several convolutions and dropouts for each model. Thus, the trial results will provide important information regarding the CNN approach which has better performance and the most appropriate number of convolutions to be used so that the stroke classification performance can be optimal.

Convolutional neural network (CNN) method was developed through several stages, i.e., building model, training model, evaluation model, and final prediction model. The CNN structure consists of two processes, feature extraction and classification [12]. The feature extraction process in CNN consists of several hidden layers. Where the hidden layer consists of a convolution layer, an activation function, and a pooling layer. CNN works hierarchically so that it can produce output in the first convolution layer. The output of the first convolution layer will be used as input in the next convolution layer. The classification process uses a fully-connected layer. Both layer components can be equipped with an activation function (softmax) for optimization of the CNN process. The input data has an image pixel size and the extraction process is carried out on the convolution layer so that its size will be reduced [13].

The convolutional layer forms a filter with length and height called the kernel. This kernel will be shifted to all parts of the input image to get important information from an object [14]. If the image is still too large, then down-sampling is done to reduce the size of the array, called max-pooling, or by taking the largest pixel value from each kernel. In the pooling process, the input image size will be drastically reduced. However, even if the number of parameters is reduced, important information from the image can still be retrieved. The process continues until the last pooling layer that produces the appropriate image size. In the classification section, there are several fully connected layers. One of them is flattened which is used for flattening data in the form of multi-dimensional arrays into one dimension. Another layer is the hidden layer and the last is the output layer, which consists of several classes that represent the output classification.

The data used is divided into training data, validation, and test data. Then, a network design was carried out using the CNN method to classify. The training data is used to carry out the network learning process and then evaluated. If the accuracy of the network model training process is not good, it is necessary to modify the CNN architecture, the network parameters, or the sample data. If the accuracy results are satisfactory, the next step is testing with data validation. The term "validation data" refers to data that is not used in the training process. If the accuracy of this validation data is poor, overfitting may occur, and the network must be modified again. If the results are good, it can be used

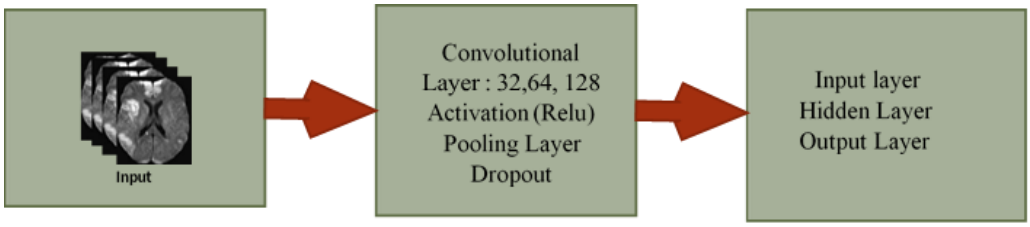

Figure. 2 CNN architecture in this study

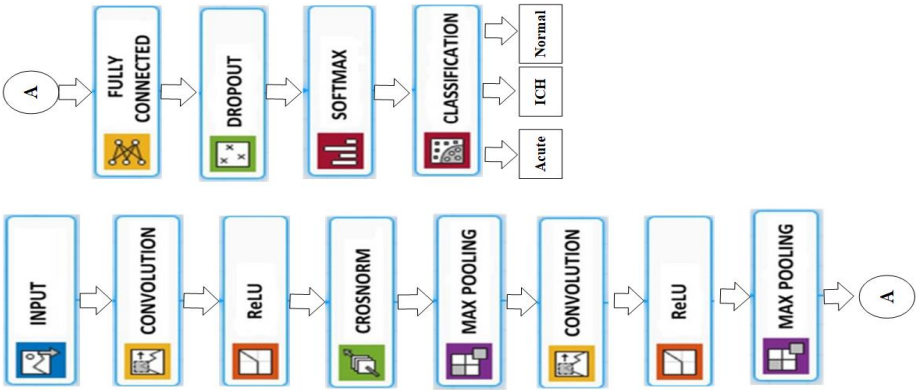

Figure. 3 CNN scenario 1st model

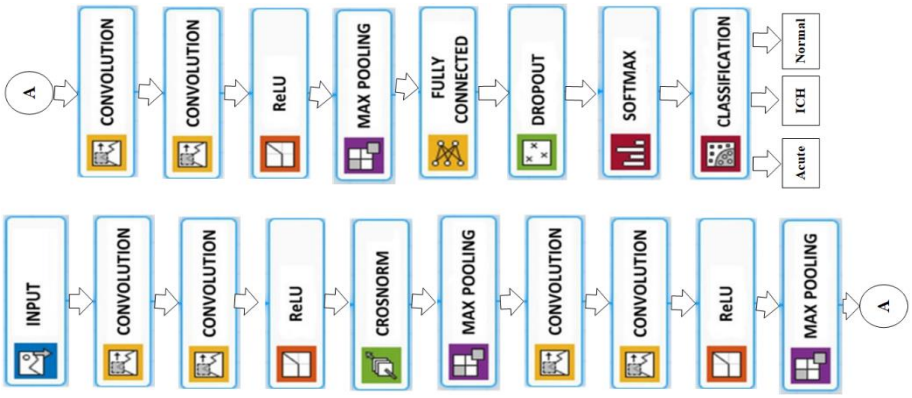

Figure. 4 CNN scenario 2nd model / HCL-NN

Table 1. DW-MRI image dataset specification

Process	Normal Images	ICH Images	Acute Ischemic Images
Training	976	976	976
Validation	418	418	418
Testing	348	348	348
Total	1,742	1,742	1,742

to process test data. The CNN architecture used to detect stroke in this study is shown in Fig. 2.

The convolutional neural network architecture used has 5 convolution layers, with dimension values 32, 64, 128. The convolutional layer is used for convolution operations on the output of the previous layer. This layer is the main activity in the convolutional neural network approach. The convolution process on the image data is carried out to extract features from the input image. This process produces a linear transformation of the input data according to the spatial information in the data.

The CNN method in this study was also modified by using ELM in the classification process, as shown in Fig. 2. CNN is used in the feature extraction process. Then, these features are entered into the ELM classifier engine. Theoretically, ELM has a fast training speed and high training precision. In addition, the use of ELM can result in better CNN performance. The dataset for the trial process in this study used a collection of stroke images (intra cerebral hemorrhage/ICH, acute ischemic stroke) and normal axial images obtained from various hospitals in Indonesia. The specifications for the number of datasets used in this study are shown in Table 1. Determination of the amount and specification of this data refers to the provisions of 80% training data/validation data and 20% test data.

Table 2. The first scenario for 1st CNN model

1st Model	Training Process				Testing Process				Optimizer
	Precision	Recall	F1-score	Support	Precision	Recall	F1-score	Support	
Acute	0.9293	0.987	0.9573	1306	0.789	0.8578	0.822	436	RMSprop
ICH	0.9861	0.925	0.9546	1306	0.7211	0.9014	0.8012	436	
Normal	1.0000	1.0000	1.0000	1307	1.0000	0.6621	0.7967	435	
Accuracy			0.9707	3919			0.8072	1307	
Macro avg	0.9718	0.9706	0.9706	3919	0.8367	0.8071	0.8066	1307	
Weighted avg	0.9718	0.9707	0.9706	3919	0.8366	0.8072	0.8066	1307	
Acute	0.9293	0.9870	0.9573	1306	0.8157	0.9037	0.8575	436	**Adam**
ICH	0.9861	0.9250	0.9546	1306	0.8964	0.7936	0.8418	436	
Normal	1.0000	1.0000	1.0000	1307	0.9932	1.0000	0.9966	435	
Accuracy			**0.9707**	3919			**0.8990**	1307	
Macro avg	0.9718	0.9706	0.9706	3919	0.9018	0.8991	0.8986	1307	
Weighted avg	0.9718	0.9707	0.9706	3919	0.9017	0.8990	0.8985	1307	
Acute	0.9405	0.9801	0.9599	1306	0.8435	0.8280	0.8356	436	SGD
ICH	0.9792	0.9380	0.9582	1306	0.8345	0.8440	0.8392	436	
Normal	1.0000	1.0000	1.0000	1307	0.9863	0.9931	0.9897	435	
Accuracy			0.9727	3919			0.8883	1307	
Macro avg	0.9732	0.9727	0.9727	3919	0.8881	0.8884	0.8882	1307	
Weighted avg	0.9732	0.9727	0.9727	3919	0.8880	0.8883	0.8881	1307	

The DWI stroke MRI image dataset has 5,226 images, consisting of 1,742 normal images, 1,742 ICH images, and 1,742 acute ischemic images, which are grouped into 3 parts, training, validation, and testing.

2.3 Testing scenarios

The trial process in this study used an MRI-DWI image dataset. The results obtained were stroke classification performance in the training, validation, and testing processes. Recognition accuracy in percent is used to measure classification performance. The original image in the dataset uses 90x90x3 pixel dimensions. The classification process has two scenarios. The first scenario uses a common CNN module which is used with 2 layers on its convolution as shown in Fig. 3.

The second scenario uses *hepta convolution layers* (HCL) as shown in Fig. 4. Each scenario uses 3 optimizers using optimizer: adam, SGD, and RMSProp. The testing process is carried out in each scenario to produce the performance of the classification model, including accuracy, precision, recall, and F1 score. Accuracy is calculated based on the ratio of the predicted data according to its class from the entire data. Precision is calculated based on the ratio of class-appropriate data compared to positive predicted data. The recall is calculated based on the ratio of predicted data according to its class compared to data that is following its class.

2.4 Performance Measure

Evaluation of classification performance, accuracy, precision, recall, and the value of the F1 score were measured. Dependent patients and independent patients who were correctly classified were considered as true positive (TP) and true negative (TN). False-positive (FP) can be interpreted as an independent patient who is misclassified as a dependent patient. False negatives (FN) are dependent patients who are misclassified as independent patients [15]. The calculation for this is as follows:

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

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (3)$$

$$\text{F1\_score} = \frac{2*\text{TP}}{2*\text{TP+FP+FN}} \quad (4)$$

3. Result and discussion

3.1 Test process with CNN model scenario

The trial process that has been carried out has resulted in the stroke classification performance as shown in Tables 2 and 3. Tables 2 and 3 show the performance of the CNN method in each test scenario determined based on the number of convolutions in

Table 3. The second scenario for model 2nd CNN model (HCL-CNN)

2nd Model	Training Process				Testing Process				Optimizer
	Precision	Recall	F1-score	Support	Precision	Recall	F1-score	Support	
Acute	1.0000	0.9770	0.9884	1306	0.9763	0.9770	0.9604	436	RMSprop
ICH	0.9775	1.0000	0.9886	1306	0.9467	1.0000	0.9616	436	
Normal	1.0000	1.0000	1.0000	1307	1.0000	1.0000	1.0000	435	
Accuracy			0.9923	3919			0.9740	1307	
Macro avg	0.9925	0.9923	0.9923	3919	0.9743	0.9740	0.9740	1307	
Weighted avg	0.9925	0.9923	0.9923	3919	0.9743	0.9740	0.9740	1307	
Acute	0.9992	0.987	0.9931	1306	0.8435	0.828	0.8356	436	Adam
ICH	0.9871	0.9992	0.9932	1306	0.8345	0.844	0.8392	436	
Normal	1.0000	1.0000	1.0000	1307	0.9863	0.9931	0.9897	435	
Accuracy			0.9954	3919			0.8883	1307	
Macro avg	0.9955	0.9954	0.9954	3919	0.8881	0.8884	0.8882	1307	
Weighted avg	0.9955	0.9954	0.9954	3919	0.888	0.8883	0.8881	1307	
Acute	0.8559	0.9502	0.9006	1306	0.8361	0.9243	0.8780	436	SGD
ICH	0.9601	0.8292	0.8899	1306	0.9465	0.8119	0.8741	436	
Normal	0.9746	1.0000	0.9872	1307	0.9645	1.0000	0.9819	435	
Accuracy			0.9265	3919			0.9120	1307	
Macro avg	0.9302	0.9265	0.9259	3919	0.9157	0.9121	0.9113	1307	
Weighted avg	0.9302	0.9265	0.9259	3919	0.9120	0.9120	0.9113	1307	

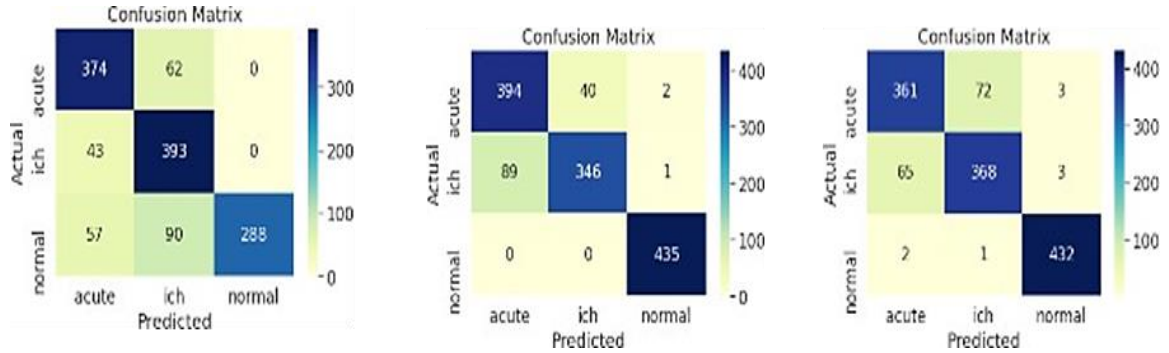

Figure. 5 Results of confusion matrix for the three optimizer models on the 1st CNN model: (a) RMSprop, (b) Adam and (c) SGD

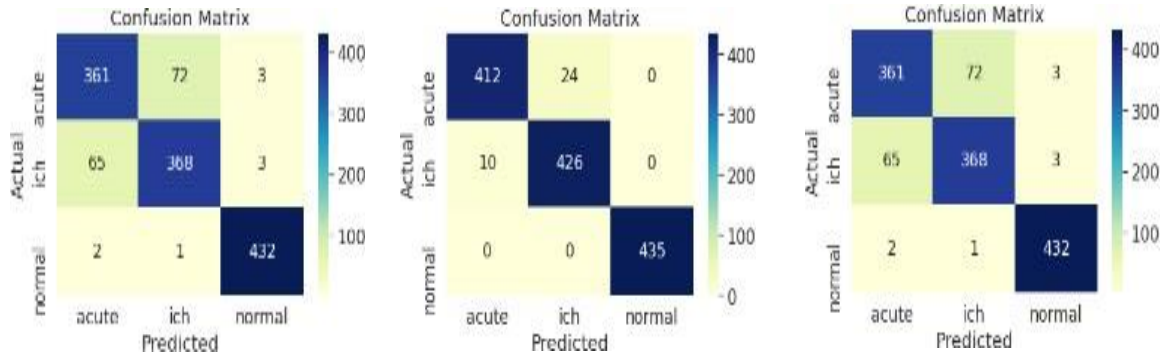

Figure. 6 Results of confusion matrix for the three optimizer models on the 2nd CNN model: (a)RMSprop, (b) Adam, and (c) SGD

each CNN layer. Based on the 3x3 confusion matrix in Fig. 5 (a), it is shown that the columns are predictions and therefore the rows must be actual values. The main diagonals (374, 393, 288) give the correct prediction. This is the case where the actual and predicted values of the model are the same. The first line is the actual acute number. The model predicted 374 of these correctly and incorrectly predicted 62 of acute becoming ICH. The second line is the actual number of ICH. The model predicted 393 of these correctly and incorrectly predicted 43 ICH to be acute. The third line is the actual normal number. The model predicted 288 of these correctly and incorrectly predicted 57 normal to acute and 90 normal to ICH.

Fig. 5 (b) shows columns are predictions and therefore rows must be actual values. The main diagonals (394, 346, 432) give the correct prediction. This is the case where the actual and predicted values of the model are the same. The first line is the actual acute number. The model predicted 394 of these correctly and incorrectly predicted 89 acute to be ICH and 2 acute to be normal. The second line is the actual number of ICH. The model predicted 346 of these correctly and incorrectly predicted 43 ICH to be acute and 1 ICH to be normal. The third line is the actual normal number. The model predicted 445 of these correctly.

Fig. 5 (c) shows columns are predictions and therefore rows must be actual values. The main diagonals (361, 368, 445) give the correct prediction. This is the case where the actual and predicted values of the model are the same. The first line is the actual acute number. The model predicted 361 of these correctly and incorrectly predicted 72 acute to be ICH and 3 acute to be normal. The second line is the actual number of ICH. The model predicted 368 of these correctly and incorrectly predicted 65 ICH to be acute and 3 ICH to be normal. The third line is the actual normal number. The model predicted 432 of these correctly and incorrectly predicted 2 normal to be acute and 1 normal to be ICH. In Table 5.1 (scenario model 1), the highest value accuracy in the testing process on the Adam optimizer is 89.90 percent compared to the RMSprop and SGD optimizers. A total of 89 ischemic stroke patients with ICH were detected as acute stroke patients, 40 acute ischemic stroke patients were detected as ICH stroke patients. 1 patient ICH stroke images were detected as normal images and 2 acute ischemic stroke images were detected as normal images.

Based on the 3x3 confusion matrix in Fig. 6 (a) it is shown that the columns are predictions and therefore the rows must be actual values. The main diagonals (361, 368, 432) give the correct prediction. This is the case where the model's actual and predicted values are the same. The first line is the actual acute number. The model predicted 361 of these correctly and incorrectly predicted 72 acute to be ICH and 3 acute to be normal. The second line is the actual number of ICH. The model predicted 368 of these correctly and incorrectly predicted 65 ICH to be acute and 3 ICH to be normal. The third line is the actual normal number. The model predicted 432 of these correctly and incorrectly predicted 2 normal to be acute and 1 normal to be ICH.

Fig. 6 (b) shows columns are predictions and therefore rows must be actual values. The main diagonals (394, 346, 432) give the correct prediction. This is the case where the model's actual and predicted values are the same. The first line is the actual acute number. The model predicted 394 of these correctly and incorrectly predicted 89 acute to be ICH and 2 acute to be normal. The second line is the actual number of ICH. The model predicted 346 of these correctly and incorrectly predicted 43 ICH to be acute and 1 ICH to be normal. The third line is the actual normal number. The model predicted 445 of these correctly.

Fig. 6 (c) shows columns are predictions and therefore rows must be actual values. The main diagonals (361, 368, 432) give the correct prediction. This is the case where the model's actual and predicted values are the same. The first line is the actual acute number. The model predicted 361 of these correctly and incorrectly predicted 72 acute to be ICH and 3 acute to be normal. The second line is the actual number of ICH. The model predicted 368 of these correctly and incorrectly predicted 65 ICH to be acute and 3 ICH to be normal. The third line is the actual normal number. The model predicted 432 of these correctly and incorrectly predicted 2 normal to be acute and 1 normal to be. Table 5.2 shows the accuracy of the testing process for the 3 classes in the stroke classification generated on the RMSprop optimizer of 97.40 percent. Recall, precision and F1-score also have the highest scores on other optimizers such as Adam, and SGD. The process of adding convolutions for each layer results in the performance value of the CNN model adding to the accuracy value and other values. In the confusion matrix RMSprop optimizer, 2 images of ischemic stroke patients were detected as normal images, 65 images of acute ischemic stroke which were read as ICH stroke images, 1 image of ICH stroke which was detected as normal images, and 72 images of acute ischemic stroke which were read as ICH strokes. On the other hand, 3 images of acute ischemic stroke were read as normal images and 3 images of ICH stroke were detected as normal images.

3.2 Trial process with transfer learning

3.2.1. ResNet50

ResNet-50 is an architecture from CNN which introduces a new concept, i.e., shortcut connections. The emergence of the shortcut connection concept in the ResNet-50 architecture is caused by the vanishing gradient problem. Vanishing gradient occurs when the process of deepening the network structure is carried out [16]. However, deepening the network with the aim of improving its performance cannot be done simply by stacking layers. This is because, the deeper the network, the more likely it is to experience

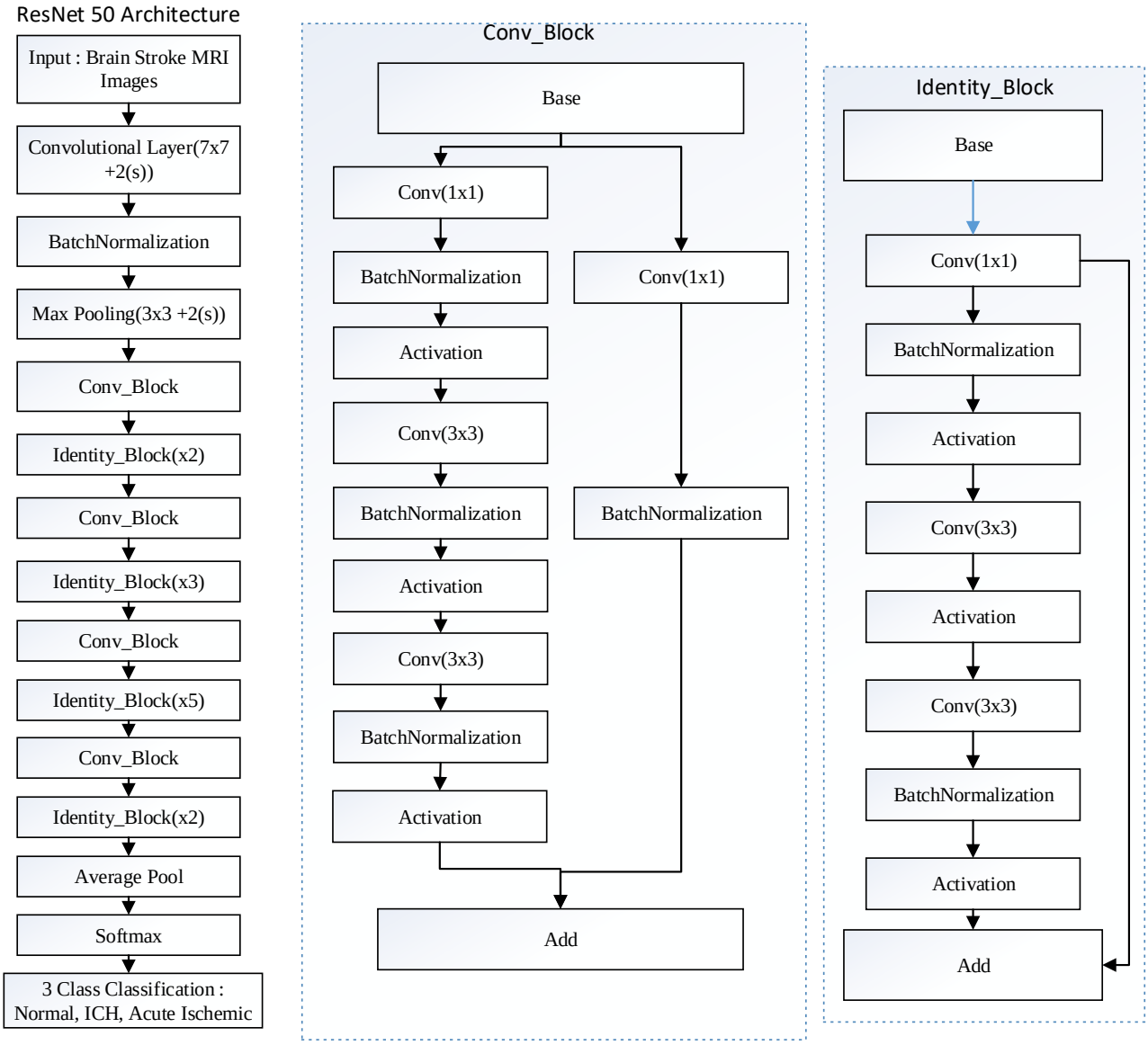

Figure. 7 ResNet-50 architecture block diagram

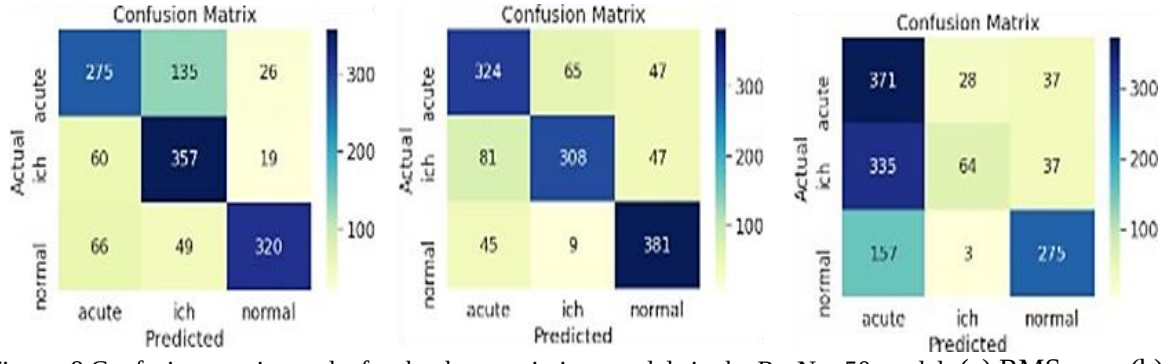

Figure. 8 Confusion matrix results for the three optimizer models in the ResNet-50 model: (a) RMSprop, (b) Adam and (c) SGD

missing gradient problems, which causes the gradient to become very small, resulting in decreased performance or accuracy.

Fig. 7 shows the difference between the ResNet-50 architecture used in this study and the original architecture. This is because there are slight modifications made to suit the needs of the classification system. Modifications made include:

Table 4. ResNet-50 architecture performance

ResNet 50 Model	Training Process				Testing Process				Optimizer
	Precision	Recall	F1-score	Support	Precision	Recall	F1-score	Support	
Acute	0.7206	0.6792	0.6993	1306	0.6858	0.6307	0.6571	436	
ICH	0.7008	0.8231	0.7570	1306	0.6599	0.8188	0.7308	436	
Normal	0.9012	0.7957	0.8452	1307	0.8767	0.7356	0.8000	435	
Accuracy			0.7660	3919			0.7284	1307	RMSprop
Macro avg	0.7742	0.7660	0.7672	3919	0.7408	0.7284	0.7293	1307	
Weighted avg	0.7742	0.7660	0.7672	3919	0.7407	0.7284	0.7293	1307	
Acute	0.7231	0.7519	0.7372	1306	0.7200	0.7431	0.7314	436	
ICH	0.8356	0.7159	0.7711	1306	0.8063	0.7064	0.7531	436	
Normal	0.8135	0.8975	0.8534	1307	0.8021	0.8759	0.8374	435	Adam
Accuracy			0.7885	3919			0.7751	1307	
Macro avg	0.7907	0.7884	0.7873	3919	0.7761	0.7751	0.7739	1307	
Weighted avg	0.7907	0.7885	0.7873	3919	0.7761	0.7751	0.7739	1307	
Acute	0.4400	0.8568	0.5814	1306	0.4299	0.8509	0.5712	436	
ICH	0.7393	0.1455	0.2431	1306	0.6737	0.1468	0.2411	436	
Normal	0.7730	0.6618	0.7131	1307	0.7880	0.6322	0.7015	435	SGD
Accuracy			0.5547	3919			0.5432	1307	
Macro avg	0.6508	0.5547	0.5126	3919	0.6305	0.5433	0.5046	1307	
Weighted avg	0.6508	0.5547	0.5126	3919	0.6304	0.5432	0.5044	1307	

a. Changing the number of outputs on the fully connected layer to 3 classes.
b. Changing the activation function on the fully connected layer from softmax to sigmoid.
c. Overriding optimization functions: RMSprop, Adam, and SGD

Based on the 3x3 confusion matrix in Fig. 8 (a), it is shown that the columns are predictions and therefore the rows must be actual values. The main diagonal (275, 357, 320) gives the correct prediction. This is the case where the actual and predicted values of the model are the same. The first line is the actual acute number. The model predicted 275 of these correctly and incorrectly predicted 135 acute to be ICH and 26 acute to be normal. The second line is the actual number of ICH. The model predicted 357 of these correctly and incorrectly predicted 47 ICH to be acute and 19 ICH to be normal. The third line is the actual normal number. The model predicted 320 of these correctly and incorrectly predicted 66 normal to acute and 49 normal to ICH.

Fig. 8 (b) indicates column and row predictions are actual values. The main diagonal (324, 308, 381) gives the correct predictive value. This is the case where the actual value and the predicted value of the model are the same. The first line is the actual acute number. The model predicted 324 of these correctly and incorrectly predicted 65 acute to be ICH and 47 acute to be normal. The second line is the actual number of ICH. The model predicted 308 of these correctly and incorrectly predicted 81 ICH to be acute and 64 ICH to be normal. The third line is the actual normal number. The model predicted 381 of these correctly and incorrectly predicted 45 normal to be acute and 9 normal to be ICH.

Fig. 8 (c) shows columns are predictions and therefore rows must be actual values. The main diagonals (371, 64, 275) give the correct prediction. This is the case where the actual and predicted values of the model are the same. The first line is the actual acute number. The model predicted 371 of these correctly and incorrectly predicted 28 acute to be ICH and 37 acute to be normal. The second line is the actual number of ICH. The model predicted 64 of these correctly and incorrectly predicted 335 ICH to be acute and 3 ICH to be normal. The third line is the actual normal number. The model predicted 275 of these correctly and incorrectly predicted 157 normal to acute and 3 normal to ICH. Table 4 shows that Adam's optimization function resulted in training accuracy of 0.7885 and testing accuracy of 0.7751. In the confusion matrix, there were 81 ICH images detected by acute ischemic stroke, 45 normal images detected by acute ischemic stroke, 9 normal images detected as acute ischemic images. 28 acute ischemic images were detected as ICH images, 47 acute ischemic images were detected as normal images. 47 ICH images were detected as normal images.

3.2.2. VGG19

The VGG19 architecture is a modification of the VGG16 architecture. This architecture has a total of 19 layers consisting of 16 convolution layers and 3 fully connected layers. The difference between this architecture and the VGG16 is only in the third, fourth and fifth parts[17]. Each section is simply

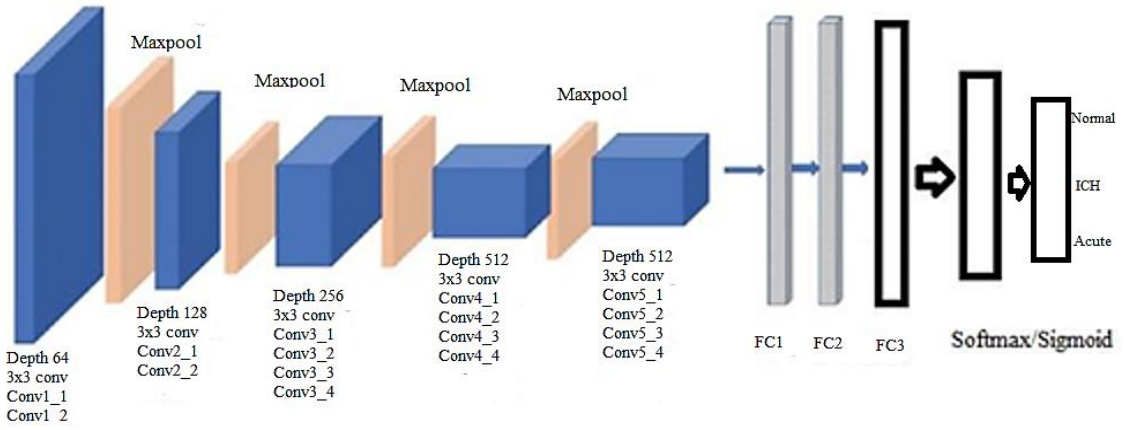

Figure. 9 VGG19 architecture

Table 5. VGG19 architecture performance

VGG19	Training Process				Testing Process				Optimize
	Precision	Recall	F1-score	Support	Precision	Recall	F1-score	Support	
Acute	0.9106	0.9900	0.9486	1306	0.8873	0.9748	0.9290	436	RMSprop
ICH	0.9891	0.9020	0.9435	1306	0.9770	0.8784	0.9251	436	
Normal	0.9992	1.0000	0.9996	1307	0.9954	0.9977	0.9966	435	
Accuracy			0.9640	3919			0.9503	1307	
Macro avg	0.9663	0.9640	0.9639	3919	0.9532	0.9503	0.9502	1307	
Weighted avg	0.9663	0.9640	0.9639	3919	0.9532	0.9503	0.9502	1307	
Acute	0.9569	0.9862	0.9713	1306	0.9369	0.9541	0.9455	436	**Adam**
ICH	0.9858	0.9564	0.9709	1306	0.9492	0.9427	0.9459	436	
Normal	1.0000	0.9992	0.9996	1307	1.0000	0.9885	0.9942	435	
Accuracy			**0.9806**	3919			**0.9617**	1307	
Macro avg	0.9809	0.9806	0.9806	3919	0.9620	0.9618	0.9619	1307	
Weighted avg	0.9809	0.9806	0.9806	3919	0.9620	0.9617	0.9618	1307	
Acute	0.8902	0.7634	0.8219	1306	0.8699	0.7362	0.7975	436	SGD
ICH	0.7934	0.9204	0.8522	1306	0.7689	0.9083	0.8328	436	
Normal	0.9720	0.9549	0.9633	1307	0.9764	0.9494	0.9627	435	
Accuracy			0.8796	3919			0.8646	1307	

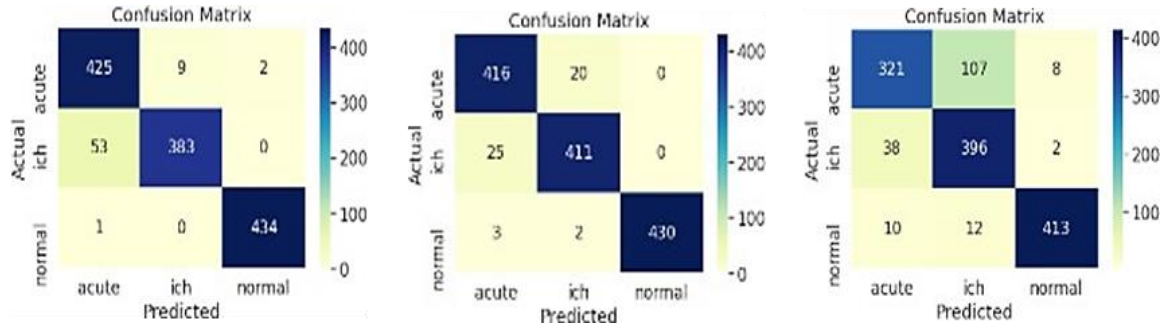

Figure. 10 Results of confusion matrix for the three optimizer models in the VGG19 model: (a) RMSprop, (b) Adam and (c) SGD

added with a convolution layer with the same configuration for each section, as shown in Fig. 9. Modifications made include:

- Changing the number of outputs in the fully connected layer to 3 classes.
- Changing the activation function on the fully connected layer from softmax to sigmoid.
- Overriding optimization functions: RMSProp, Adam and SGD.

From the experiment produced on Adam's optimization function, the training accuracy is 0.9806 and the test accuracy is 0.9617. 25 ICH images were detected as acute ischemic images, 3 normal images were detected as acute ischemic images. 20 acute ischemic images were detected as ICH images, 2 normal images were detected as ICH images, as shown in Table 9.

Based on the 3x3 confusion matrix in Fig. 10 (a), it is shown that the columns are predictions and therefore the rows must be actual values. The main diagonals (425, 383, 434) give the correct prediction. This is the case where the actual and predicted values of the model are the same. The first line is the actual acute number. The model predicted 425 of these correctly and incorrectly predicted 9 acute to be ICH and 2 acute to be normal. The second line is the actual number of ICH. The model predicted 383 of these correctly and incorrectly predicted 53 ICH to be acute. The third line is the actual normal number. The model predicted 434 of these correctly and incorrectly predicted 1 normal to be acute.

Fig. 10 (b) shows columns are predictions and therefore rows must be actual values. The main diagonal (416, 411, 430) gives the correct prediction. This is the case where the actual and predicted values of the model are the same. The first line is the actual acute number. The model predicted 416 of these correctly and incorrectly predicted 20 acute to be ICH. The second line is the actual number of ICH. The model predicted 411 of these correctly and incorrectly predicted 25 ICH to be acute. The third line is the actual normal number. The model predicted 430 of these correctly, and incorrectly predicted 3 normal to be acute and 2 normal to be ICH.

Fig. 10 (c) shows columns are predictions and therefore rows must be actual values. The main diagonals (321, 396, 413) give the correct prediction. This is the case where the actual and predicted values of the model are the same. The first line is the actual acute number. The model predicted 321 of these correctly and incorrectly predicted 107 acute to be ICH and 8 acute to be normal. The second line is the actual number of ICH. The model predicted 64 of these correctly and incorrectly predicted 335 ICH to be acute and 3 ICH to be normal. The third line is the actual normal number. The model predicted 413 of these correctly and incorrectly predicted 10 normal to be acute and 12 normal to be ICH.

3.2.3. MobileNetV2

MobileNetV2 still uses depthwise and pointwise convolution. MobileNetV2 adds two new features: 1) linear bottlenecks, and 2) shortcut connections between bottlenecks[18]. The basic structure of this architecture is shown in Fig. 11.

The bottleneck contains the inputs and outputs between the models, while the inner layer encapsulates the model's ability to convert inputs from lower-level concepts (i.e., pixels) to higher-level descriptors. Ultimately, as with residual connections on traditional CNNs, shortcuts between bottlenecks allow for faster training and better accuracy. From the experiment, the optimization function RMSprop resulted in a training accuracy of 0.9575 and a testing accuracy of 0.9419. 51 ICH images were detected as acute images. 22 acute ischemic images were detected as ICH images. 3 acute ischemic images were detected as normal images. The confusion matrix is attached in Table 6.

Based on the 3x3 confusion matrix in Fig. 12 (a), it is shown that the columns are predictions and, therefore, the rows must be actual values. The main diagonals (411, 385, 435) give the correct prediction. With the problem where the actual and predicted values of the model are the same. The first line is the actual acute number. The model predicted 411 of these correctly and incorrectly predicted 22 acute ICH and three acute to be expected. The second line is the actual number of ICH. The model predicted 385 of these correctly and incorrectly predicted 51 ICH to be acute. The third line is the actual standard number. The model predicted 435 of these correctly.

Fig. 12 (b) shows columns are predictions, and therefore rows must be actual values. The main diagonals (403, 354, 431) are correctly predicted. This is the case where the actual and predicted values of the model are the same. The first line is the actual acute number. The model predicted 403 of these correctly and incorrectly predicted 20 acute to be ICH and 13 acute to be expected. The second line is the actual number of ICH. The model predicted 354 of these correctly and incorrectly predicted 79 ICH to be acute and 3 ICH to be expected. The third line is the actual standard number. The model predicted 435 of these correctly.

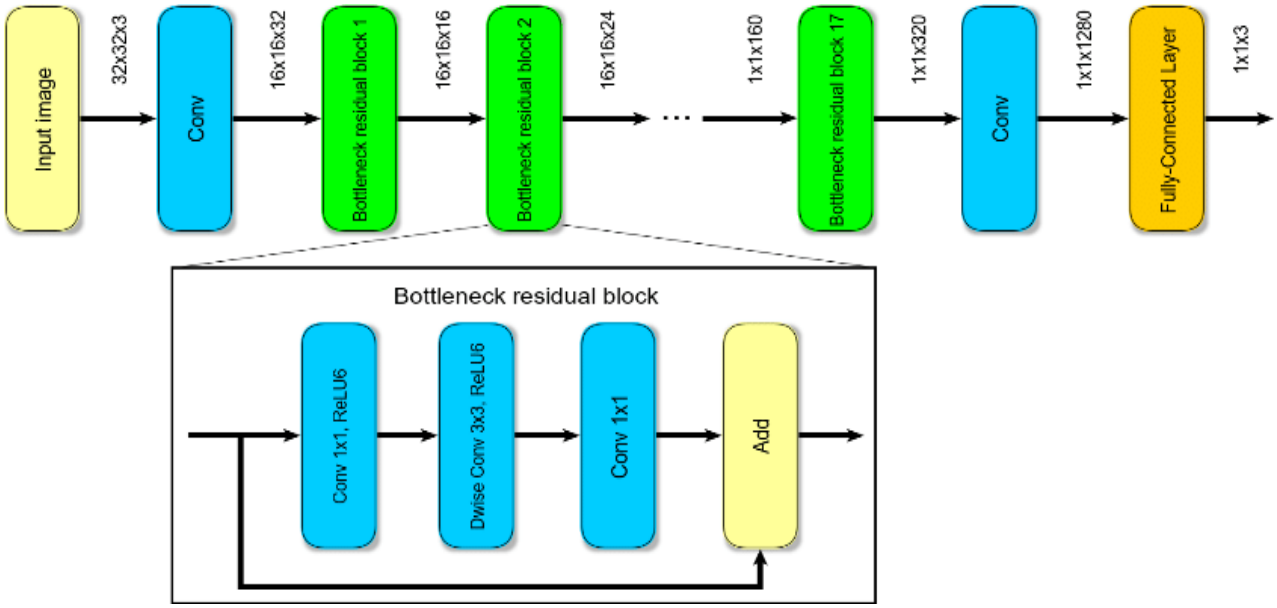

Figure. 11 Architecture of mobilenetV2

Table 6. CNN-mobilenetV2 performance

MobileNetV2	Training Process				Testing Process				Optimizer
	Precision	Recall	F1-score	Support	Precision	Recall	F1-score	Support	
Acute	0.9101	0.9686	0.9384	1306	0.8896	0.9427	0.9154	436	
ICH	0.9664	0.9035	0.9339	1306	0.9459	0.8830	0.9134	436	
Normal	0.9992	1.0000	0.9996	1307	0.9932	1.0000	0.9966	435	
Accuracy			0.9574	3919			0.9419	1307	RMSprop
Macro avg	0.9586	0.9574	0.9573	3919	0.9429	0.9419	0.9418	1307	
Weighted avg	0.9586	0.9574	0.9573	3919	0.9429	0.9419	0.9417	1307	
Acute	0.8559	0.9502	0.9006	1306	0.8361	0.9243	0.8780	436	
ICH	0.9601	0.8292	0.8899	1306	0.9465	0.8119	0.8741	436	
Normal	0.9746	1.0000	0.9872	1307	0.9645	1.0000	0.9819	435	Adam
Accuracy			0.9265	3919			0.9120	1307	
Macro avg	0.9302	0.9265	0.9259	3919	0.9157	0.9121	0.9113	1307	
Weighted avg	0.9302	0.9265	0.9259	3919	0.9157	0.9120	0.9113	1307	
Acute	0.9495	0.9364	0.9429	1306	0.9416	0.8876	0.9138	436	
ICH	0.9367	0.9518	0.9442	1306	0.8937	0.9450	0.9186	436	
Normal	1.0000	0.9977	0.9989	1307	0.9908	0.9908	0.9908	435	SGD
Accuracy			0.9620	3919			0.9411	1307	
Macro avg	0.9621	0.9620	0.9620	3919	0.9420	0.9411	0.9411	1307	
Weighted avg	0.9621	0.9620	0.9620	3919	0.9420	0.9411	0.9410	1307	

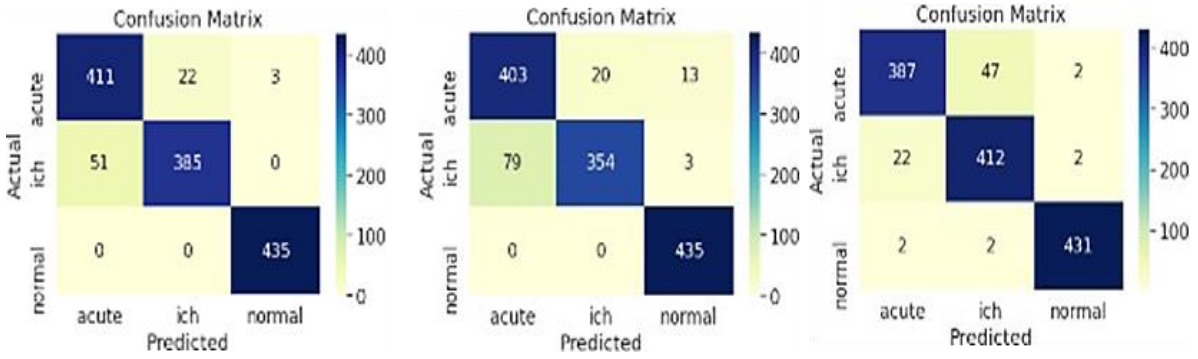

Figure. 12 Results of confusion matrix for the three optimizer models in the mobilenetV2 model: (a) RMSprop, (b) Adam and (c) SGD

Table 7. Experiment results with DW-MRI researcher data (n = 5,266; optimizer: RMSprop)

Prior Work	Network	Number of Convolution Layers	ACC (%)
Gautam et al. [19]	Multi-Focus Image Fusion and CNN	13	Training: 86.076; Testing : 81.389
Do et al. [20]	The Recurrent Residual Convolutional Neural Network (RRCNN)	16	Training = 98.090;Testing = 97.222
Pareira et al. [21]	2 step Convolution Layer		
	1st Step	5	Training: 94.375; Testing = 90.833
	2nd Step	3	Training: 97.986; Testing = 93.750
HCL-NN	***Hepta Convolutional Layer Neural Network***	**7**	**Training = 99.230; Testing = 97.40**

Fig. 12 (c) shows columns are predictions and therefore rows must be actual values. The main diagonals (387, 412, 413) give the correct prediction. With the actual and predicted values of the model are the same, the first line is the actual acute number. The model predicted 387 of these correctly and incorrectly predicted 47 acute to be ICH and 2 acute to be normal. The second line is the actual number of ICH. The model predicted 412 of these correctly and incorrectly predicted 22 ICH to be acute and 2 ICH to be normal. The third line is the actual normal number. The model predicted 431 of these correctly and incorrectly predicted 2 normal to be acute and 2 normal to ICH.

The proposed method is also compared with other researchers. Gautam et al. [19] proposed a method using image fusion and CNN approaches. Initially, several preprocessing operations have been used using multi-focus image fusion to improve image quality. Next, the preprocessed image is entered into the CNN architecture with 13 layers of stroke classification. Do et al. [20] developed a deep learning algorithm for the automated binary classification of the alberta stroke program early computed tomographic score (ASPECTS) using imaging diffusion weighting (DWI) in acute stroke patients. The recurrent residual convolutional neural network (RRCNN) was developed for the classification of DWI stroke images. CNN uses 16 convolution layers. Pareira et al. [21] have classified stroke using 2 CNN approaches. The first approach uses 5 convolution layers and the second approach uses 3 convolution layers.

The dataset used in the study, with a total of 5,266 images (1,742 every class) and used RMSprop optimizer. Then, the data was tested on the convolution layer model from the previous researchers (Table 7). The results of these experiments are written in Table 7. Gautam et al. [19] adopted the VGG16 and resnet, employing 13 convolution layers, the accuracy training result shows 86.076 percent and accuracy testing 81.389 percent. As for Do et al. [20], which employs sixteen convolution layers, the accuracy training result shows 98.090 percent and accuracy testing 97.222 percent. The research conducted by H. Pareira et al. [21] adopted 5 convolution layers for 1st step, the accuracy training result shows 94.375 percent and accuracy testing 90.833 percent. And the 2nd step, it adopted 3 convolution layers, the accuracy training result shows 97.986 percent and accuracy testing 93.750 percent. In this study, we used seven convolution layers in the proposed ***hepta convolutional layer neural network* (HCL-NN)**, and the experiment deliver the highest accuracy training value **99.230** percent and the highest accuracy testing value **97.40** percent.

Conclusion

The experiment in this study shows by using scenario hepta convolution layer neural network (HCL-NN) display data training accuracy value 99.23 percent and data testing accuracy 97.40 percent. This value is higher compare to the original scenario CNN. This study was running RMSprop optimizer to present those both values.

This study also tested the values of precision, recall and F1-scores. The scores for testing the HCL-NN scenario 97.43, 97.40 and 97.40 are higher compare to the other scenarios and other transfer learning

As summary this study that HCL-NN showed the best scenario CNN model for MR-DWI images

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.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

Conceptualization: Andi Kurniawan Nugroho; Methodology: Andi Kurniawan Nugroho; Software: Andi Kurniawan Nugroho; Validation: Terawan Agus Putranto; Formal analysis: Andi Kurniawan Nugroho; Writing original draft preparation: Andi Kurniawan Nugroho and Dinar M K Nugraheni; Writing review and editing: Terawan Agus Putranto, I Ketut Eddy Purnama, Dinar M K Nugraheni and Mauridhi Hery Purnomo;Visualization: Mauridhi Hery Purnomo; Supervision: Mauridhi Hery Purnomo;Funding acquisition: I Ketut Eddy Purnama.

References

[1] D. Kuriakose, "Pathophysiology and Treatment of Stroke : Present Status and Future Perspectives", *Int. J. Mol. Sci.*, Vol. 21, pp. 1–24, 2020, doi: 10.3390/ijms21207609.

[2] I. Setyopranoto, H. F. Bayuangga, A. S. Panggabean, S. Alifaningdyah, L. Lazuardi, F. S. T. Dewi, and R. G. Malueka, "Prevalence of Stroke and Associated Risk Factors in Sleman District of Yogyakarta Special Region, Indonesia", *Stroke Research and Treatment*, Vol. 2019, p. 2642458, May 2019, doi: 10.1155/2019/2642458.

[3] L. D. Mccullough, "Age and Sex Are Critical Factors in Ischemic Stroke Pathology", *Endocrinology*, Vol. 159, No. May, pp. 3120–3131, 2018, doi: 10.1210/en.2018-00465.

[4] J. Lu, Q. Mei, X. Hou, A. Manaenko, L. Zhou, D. S. Liebeskind, J. H. Zhang, Y. Li, and Q. Hu, "Imaging Acute Stroke: From One-Size-Fit-All to Biomarkers", *Frontiers in Neurology*, Vol. 12, 2021, doi: 10.3389/fneur.2021.697779.

[5] S. H. You, B. Kim, B. K. Kim, and S. E. Park, "Fast MRI in Acute Ischemic Stroke: Applications of MRI Acceleration Techniques for MR-Based Comprehensive Stroke Imaging", *Investigative Magnetic Resonance Imaging., Korean Society of Magnetic Resonance in Medicine*, Vol. 25, No. 2, p. 81, 2021.

[6] O. D. Nurhayati, "Stroke Identification System on the Mobile Based CT Scan Image", In: *Proc. of Conference on Information Technology, Computer and Electrical Engineering (ICITACEE)*, pp. 113–116, 2015.

[7] M. Malathi and P. Sinthia, "Brain Tumour Segmentation Using Convolutional Neural Network with Tensor Flow", *Asian Pacific J. Cancer Prev.*, Vol. 20, pp. 2095–2101, 2019, doi: 10.31557/APJCP.2019.20.7.2095.

[8] P. Rao, N. A. Pereira, and R. Srinivasan, "Convolutional Neural Networks for Lung Cancer Screening in Computed Tomography (CT) Scans", In: *Proc. of 2nd International Conference on Contemporary Computing and Informatics (ic3i)*, 2016, pp. 489–493.

[9] S. Sarraf, D. D. Desouza, and J. Anderson, "DeepAD : Alzheimer ' s Disease Classification via Deep Convolutional Neural Networks using MRI and fMRI", pp. 1–32, 2017.

[10] F. Wahid, R. Ghazali, U. Tun, H. Onn, and A. S. Shah, "Using Probabilistic Classification Technique and Statistical Features for Brain Magnetic Resonance Imaging (MRI) Classification: An Application of AI Technique in Bio-Science Using Probabilistic Classification Technique and Statistical Features for Bra", *Int. J. Bio-Science Bio-technology*, Vol. 8, No. 6, pp. 93–106, 2016, doi: 10.14257/ijbsbt.2016.8.6.10.

[11] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning", *J. Big Data*, Vol. 6, No. 60, pp. 1–48, 2019, doi: 10.1186/s40537-019-0197-0.

[12] R. Yamashita, M. Nishio, R. Kinh, G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology", *Insights Imaging*, Vol. 9, pp. 611–629, 2018, doi: https://doi.org/10.1007/s13244-018-0639-9.

[13] N. F. Nissa, A. Janiati, N. Cahya, and P. Astuti, "Application of Deep Learning Using Convolutional Neural Network (CNN) Method for Women's Skin Classification", *Sci. J. Informatics*, Vol. 8, No. 1, pp. 144–153, 2021, doi: 10.15294/sji.v8i1.26888.

[14] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, and B. Shuai, "Recent Advances in Convolutional Neural Networks", *Comput. Vis. Pattern Recognit.*, pp. 1–38, 2017.

[15] J. E. L. García, C. Y. Márquez, Y. V. Rey, and O. C. Nieto, "A transfer learning method for pneumonia classification and visualization",

Appl. Sci., Vol. 10, No. 8, 2020, doi: 10.3390/APP10082908.

[16] S. A. T. W. L. Passos, C. R. Pereira, E. Rezende, Tiago Carvalho, "Parkinson Disease Identification Using Residual Networks and Optimum-Path IFSP - Federal Institute of S ˜", In: *Proc. of IEEE 12th International Symposium on Applied Computational Intelligence and Informatics*, No. January 2019, doi: 10.1109/SACI.2018.8441012.

[17] X. Zhang, J. Zou, K. He, and J. Sun, "Accelerating Very Deep Convolutional Networks for Classification and Detection", *Comput. Vis. Pattern Recognit.*, pp. 1–14, 2015.

[18] M. Sandler, M. Zhu, A. Zhmoginov, and C. V Mar, "MobileNetV2: Inverted Residuals and Linear Bottlenecks", In: *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4510–4520, 2019.

[19] A. Gautam and B. Raman, "Biomedical Signal Processing and Control Towards effective classification of brain hemorrhagic and ischemic stroke using CNN", *Biomed. Signal Process. Control*, Vol. 63, No. August 2020, p. 102178, 2021, doi: 10.1016/j.bspc.2020.102178.

[20] L. N. Do, B. H. Baek, S. K. Kim, H. J. Yang, I. Park, and W. Yoon, "Automatic Assessment of ASPECTS Using Diffusion-Weighted Imaging in Acute Ischemic Stroke Using Recurrent Residual Convolutional Neural Network", *Diagnostics*, Vol. 10, No. 10, 2020, doi: 10.3390/diagnostics10100803.

[21] D. R. Pereira, P. P. R. Filho, G. H. D. Rosa, J. P. Papa, and V. H. C. D. Albuquerque, "Stroke Lesion Detection Using Convolutional Neural Networks", In: *Proc. of the International Joint Conference on Neural Networks*, Vol. 2018-July, 2018, doi: 10.1109/IJCNN.2018.8489199.