

The use of CLAHE for improving an accuracy of CNN architecture for detecting pneumonia

Elbert Alfredo Tjoa¹, I Putu Yowan Nugraha Suparta², Rita Magdalena³ and Nor Kumalasari CP⁴

Abstract. Artificial intelligence (AI) has now grown rapidly for helping many aspects of human life, one of them is for medical image processing. Currently, the world is still suffering from COVID-19 pandemic outbreak which affects more than 36 million people and it is estimated that more than 1 million death occurred as a result of this outbreak. Early detection for COVID-19 sufferers is needed to assist doctors and medical experts to determine the next medication for patients for avoiding the worsening condition which leads to death. AI-based model is can be used for assisting medical experts for detecting and classify the lung condition based on chest x-ray (CXR) patient's image accurately by using deep learning. On this paper, authors proposed the use on contrast limited adaptive histogram equalization (CLAHE) for pre-processing the medical images combined with CNN AlexNet architecture. The result of this method then compared with non-CLAHE CNN AlexNet also self-made CNN architecture. The result shows a promising result by the accuracy of CNN AlexNet architecture is 91.11%.

1 Introduction

Artificial intelligence (AI) has now grown rapidly to help human life and has been applied to many aspects of living, one of which is by utilizing deep learning in image classification, especially in the process of medical imaging [1]. AI-based models on images classification are often used in computer aided diagnosis (CAD) technology to help doctors or medical experts diagnose patients more accurately. Currently, the world is suffering from the COVID-19 pandemic outbreak which affects more than 36 million people and it is estimated that more than 1 million deaths occurred as a result of this outbreak [2], [3]. Early detection in COVID-19 sufferers is needed to assist doctors in making decisions in determining follow-up actions, to prevent worsening of the patient's condition due to lower oxygen level that leads to death [4]. The main technique in diagnosing COVID-19 is to use reverse transcription polymerase chain reaction (RT-PCR), however, chest X-ray (CXR) images can be used to perform early detection of a patient's condition with an accuracy of up to 80% [5]. CAD can be used to assist doctors in determining the condition of patients based on the results of CXR images, whether the patient is infected by COVID-19 or not. Utilizing the deep learning model, the convolutional neural network (CNN), CXR images then classified into specific categories, hence, this deep learning model is very interesting to develop.

1.1 Related Work

In previous results of research, CNN is used for classifying pneumonia disease and tuberculosis, by using CXR images, there were conducted four classification. Those result shown a promising accuracy, they are binary classification (Normal vs COVID-19), the accuracy obtained is 99.7%, three classification (Normal vs COVID-19 vs Pneumonia), obtained 95.02% and four classification (Normal vs COVID-19 vs Pneumonia vs Tuberculosis (TB)) obtained 94.53% [1]. Other research conduct by Muhamet Fatih Aslan, Muhammed Fahri Unlersen, Kadir Sabanci and Akif Durdu were using a modified AlexNet and adding flatten and bidirectional long short term memories (BiLSTM). This layer were added in order to ensure the data is transformed into one-dimensional array. On this research

best accuracy obtained is 98.70% [2]. Another method was also proposed by using long-short term memories (LSTM) to detect COVID-19 based on CT images [3]. Highest accuracy that gained is 99.37%. LSTM also combined by Recurrent Neural Network (RNN) in order to detect COVID-19 based on CXR images, it obtained 95.04% accuracy after getting robust normalization. AI also implemented in detecting COVID-19, by using a modified CNN, named AI-enables smart biomedical diagnosis system (AIRBiS) to detect dan distinguish images, whether it was COVID-19 infection or pneumonia. The best accuracy gained is 97.18% [5]. In addition, transfer learning is used to detect COVID-19 Pneumonia, viral pneumonia, and normal circumstances. In this case there were eight different architecture combined with transfer learning and image augmentation to fasten the training process and increase the accuracy. The best accuracy gained is by using DenseNet201 architecture with image augmentation with an accuracy 99.70% [6]. Another study proposed a COV-MCNet (Multi-classification network) that carries eight architectures that have been modified to detect CXR images and divide it into four different classifications, namely normal conditions, COVID-19 pneumonia, viral pneumonia, and bacterial pneumonia [7].

1.2 Convolutional Neural Network (CNN)

CNN is a deep learning algorithm which usually use to process spatial data like image processing. CNN has the dynamic ability to comprehend spatial information in a gradual, low-to-high level pattern which inspired by the workings of the human nervous system[8]. Operation of CNN will have a high complexity and necessitates a large amount of data and execution time during the training period because the form of feed forward operation, this operation is a hierarchical operation which the preceding process' output results should be used as an input in the following process. CNN operation is a complex mathematical operation that generally consists of convolutional layers, down-sampling layers (pooling), activation function, and fully connected layer.

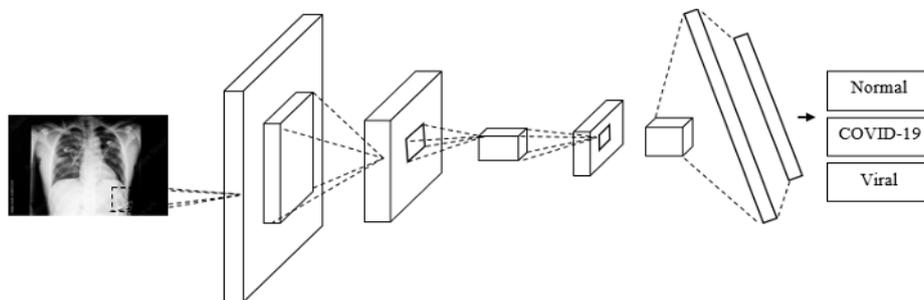


Figure 1. The Convolutional neural Network (CNN) Common Architecture

1.2.1 Convolutional Layer

The convolution layer is the main layer in the CNN process. This layer is a collection of mathematical processes that use an array of numbers to perform convolution processes that produce linear transformation processes. This process contains spatial information that can later be used for future processes on CNN [8].

1.2.2 Rel-U Activation

Rel-U is one of activation layer which makes all pixel values that are less than 0 will be made to 0, but if the pixel values are more than 0, this function will let the values remain the same [9]. Rel-U activation equation shown below :

$$f(x) = \begin{cases} 0, & x \leq 0 \\ x, & x \geq 0 \end{cases} \quad (1)$$

1.2.3 Pooling Layer

Pooling layer is a sampling layer that plays a role in reducing spatial size, lowering the amount of model parameters and also reducing the calculations complexity, so as to speed up the process of classification of data. Pooling layer is generally divided into two categories; maximum pooling which takes the highest value in each image matrix and average pooling which takes the average value of each matrix [10].

1.2.4 Fully Connected Layer

A fully activated layer has activation neurons that are interconnected to the previous and subsequent layers. Fully Connected layer will later classify the results of the image into several categories available. The results of the recognition at this layer will be in the form of percentages. In summarize, the data extraction process occurs in the convolution and pooling layer, and the extraction data is then handled at the fully connected layer to generate an output in the form of classification results [10].

1.2.5 Softmax Activation

Softmax activation layer is a activation layer which plays classification output on more than two categories. This activation layer is in form of logistic regression.

1.3 Contrast Limited Adaptive Histogram Equalization (CLAHE)

Contrast limited adaptive histogram equalization (CLAHE) is a development version of the adaptive histogram equation (AHE) that plays a role in increasing contrast in the image by increasing the intensity range of the image or performing a stretching out mechanism at the most frequent intensity value in the image [10]. In CLAHE, the image is broken down into sub-images called tiles or blocks, then performs the histogram equalization process on each sub-images that has a certain value that causes the image to be overamplified and then redistribute the pixels back to the histogram, resulting in the contrast in the image being increasingly visible [10], [11].

2 System Design

2.1 Proposed System Design

2.1.1 AlexNet

AlexNet consist of multiple sets of convolutional layers, combined with pooling layer and fully connected layer. Convolutional layers function is to extract feature of input images into pixel value, after that the process will be continued in pooling layer, where the maximum

value or average value will be used as down-sampling to make the model compact without losing any precious spatial information. After extracting features are done, fully connected layer take part to classify the result of extraction based on proposed research. The block diagram of AlexNet shown on Table 1.

Table 1. AlexNet Architecture

Layer	input	output
Input	(None, 50, 50, 3)	(None, 50, 50, 3)
Convolution 1 (Conv2D)	(None, 50, 50, 3)	(None, 13, 13, 96)
Pooling 1 (MaxPooling2D)	(None, 13, 13, 96)	(None, 7, 7, 96)
Convolution 2 (Conv2D)	(None, 7, 7, 96)	(None, 7, 7, 256)
Pooling 2 (MaxPooling2D)	(None, 7, 7, 256)	(None, 4, 4, 256)
Convolution 3 (Conv2D)	(None, 4, 4, 256)	(None, 4, 4, 384)
Convolution 4 (Conv2D)	(None, 4, 4, 384)	(None, 4, 4, 384)
Convolution 5 (Conv2D)	(None, 4, 4, 384)	(None, 4, 4, 256)
Pooling 3 (MaxPooling2D)	(None, 4, 4, 256)	(None, 2, 2, 256)
Flatten (Flatten)	(None, 2, 2, 256)	(None, 1024)
Dense (Dense)	(None, 1024)	(None, 4096)
Dense 1 (Dense)	(None, 4096)	(None, 4096)
Dense 2 (Dense)	(None, 4096)	(None, 1000)
Dense 3 (Dense)	(None, 1000)	(None, 3)
Output	(None, 3)	(None, 3)

2.1.2 YENS-Net

Layer	input	output
Input	(None, 50, 50, 3)	(None, 50, 50, 3)
Convolution 1 (Conv2D)	(None, 50, 50, 3)	(None, 50, 50, 64)
Activation 1(Activation)	(None, 50, 50, 64)	(None, 50, 50, 64)
Pooling 1 (MaxPooling2D)	(None, 50, 50, 64)	(None, 25, 25, 64)
Convolution 2 (Conv2D)	(None, 25, 25, 64)	(None, 25, 25, 32)
Activation 2(Activation)	(None, 25, 25, 32)	(None, 25, 25, 32)
Pooling 2 (MaxPooling2D)	(None, 25, 25, 32)	(None, 12, 12, 32)
Convolution 3 (Conv2D)	(None, 12, 12, 32)	(None, 12, 12, 16)
Activation 3(Activation)	(None, 12, 12, 16)	(None, 12, 12, 16)
Convolution 4 (Conv2D)	(None, 12, 12, 16)	(None, 6, 6, 16)
Flatten (Flatten)	(None, 6, 6, 16)	(None, 576)
Dense (Dense)	(None, 576)	(None, 3)
Output	(None, 3)	(None, 3)

2.1.3 Our proposed scheme

This research will be conducted to find out the most suitable parameters for each circumstance. The authors propose three scenarios, the first one is using AlexNet itself, and then the result of this architecture will be compared to CLAHE + AlexNet and YENS-Net, self-made CNN architecture. At first, the parameters, such as epoch, optimizer, batch size, and learning rate will be optimized in AlexNet architecture in order to find the best

combination for gaining the best accuracy. The same scenario is also applied to CLAHE+ AlexNet architecture and YENS-Net architecture, so every condition might have different batch size, learning rate, or even parameter.

2.2 Evaluation Parameter

2.2.1 Accuracy and Confusion Matrix

Table 2. Multi-Classes Confusion Matrix

Confusion Matrix	Actual A	Actual B	Actual C
Predicted A	X_{AA}	X_{AB}	X_{AC}
Predicted B	X_{BA}	X_{BB}	X_{BC}
Predicted C	X_{CA}	X_{CB}	X_{CC}

Confusion matrix is consist of four parts, they are :

1. True Positive (TP) is when the the actual class and predicted class shows a correct result.
2. True Negative (TP) is when the actual class and predicted class shows a false result.
3. False Positive (FP) is when the actual class is false but predicted class shows the otherwise.
4. False Negative (FN) is when the actual class is correct but the predicted class shows a false result.

Accuracy is a comparison between the data which is predicted correctly with the predicted whole data, the equation can be seen below.

$$\frac{\sum TP + \sum TN}{\sum TP + \sum TN + \sum FP + \sum FN} \quad (2)$$

n in the equation shows the category. The category for the classification is three.

3 Dataset Description

The collection of image data is done with open source data that can be accessed by public. The three types of datasets COVID-19, namely Normal and Viral Pneumonia CXRs are all obtained through the Kaggle (<https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>) site which already has the type of CXR COVID-19, Normal and pneumonia images on a large-number datasets. The total dataset used in the study was 3616 CXR COVID-19, 1345 CXR Pneumonia and 3993 Normal CXR with a total of 8954 data in format. PNG as research material. In this study, the data was divided by 75% for data training with a total data of 6715 and 25% for data testing with a total of 2239 data. Then test with test data in order to know the level of accuracy.

3.1 Pre-processing

In this research, images as input first processed by reshaping dimension of images, in order to have an uniform size of input. The size of images is resized to 50x50. Then after resizing

images, the input will be processed by AlexNet architecture also YENS architecture. After finishing the phase, the result will be compared to the images which got two step pre-processing, firstly, the images got reshaped to 50x50 and then CLAHE is applied to input images.

4 Result and Discussion

After several measurement, the best accuracy that successfully gain comes from 50x iteration (epoch) with Adamax Optimizer for every architecture (AlexNet, CLAHE + AlexNet, and YENS-Net). The result shows that there is no any significant differences between using CLAHE for pre-processing images input before it is extracted to architecture. The result can be seen in figure below.

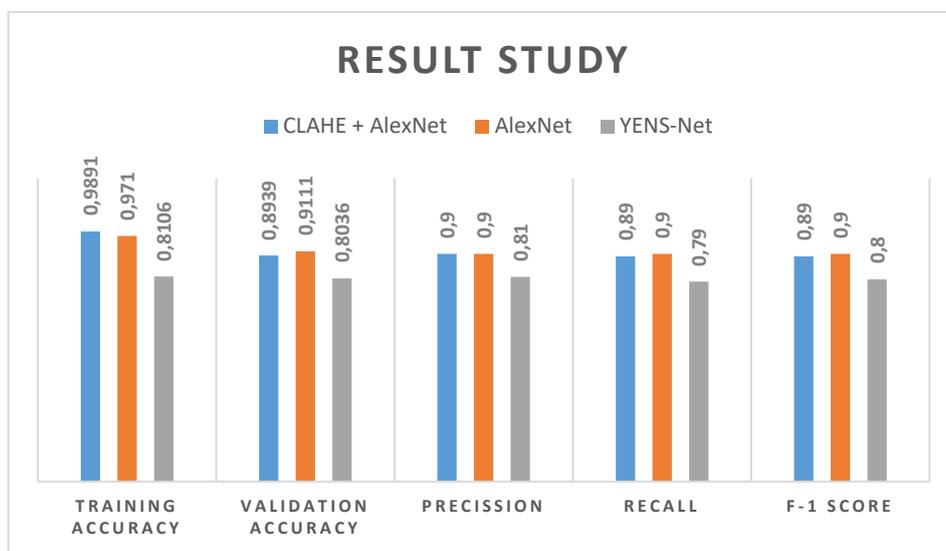


Figure 2. Result of Proposed Scenario

4.1 Best Model

After we find out the best parameters for our model, CNN AlexNet have a slightly advantages compared to CLAHE + AlexNet. The best accuracy that successfully gained is 91.11% with precision rate of 90%, recall rate is 90% and for f-1 score, it has successfully gain 90%.

5 Conclution and Future Work

This research apply two different CNN architecture with CLAHE addition in pre-processing for pneumonia detection. Initially, the authors studied the use of CLAHE for enhancing quality images for detecting and classifying pneumonia disease, authors also calculated the best parameter that might suit for each scenario and compare them, they are CNN AlexNet, CLAHE + CNN AlexNet, and YENS-Net, which is self-made CNN architecture for the same database. Finally, the result shows that there are no any significant differences the use of CLAHE for improving the accuracy. The best accuracy that successfully gained comes for AlexNet by the score 91.11%. Since COVID-19 is still raging from all over the world, the

following study can continue to detect this disease precisely, so the deterioration of patient's condition can be avoided.

References

- [1] N. N. Qaqos and O. S. Kareem, *COVID-19 Diagnosis from Chest X-ray Images Using Deep Learning Approach*, in the 3rd Int. Conf. Adv. Sci. Eng. ICOASE 2020, 110–116, (2020)
- [2] M. F. Aslan, M. F. Unlarsen, K. Sabanci, and A. Durdu, *CNN-based transfer learning–BiLSTM network: A novel approach for COVID-19 infection detection*, *Appl. Soft Comput.*, **98**, 1–12, (2021)
- [3] M. H. Memon, N. A. Golilarz, J. Li, M. Yazdi, and A. Addeh, *Early Detection of Covid-19 Disease using Computed Tomography Images and Optimized CNN-LSTM*, 2020 17th Int. Comput. Conf. Wavelet Act. Media Technol. Inf. Process. ICCWAMTIP 2020, 161–165, (2020)
- [4] S. Goyal and R. Singh, *Detection and classification of lung diseases for pneumonia and Covid-19 using machine and deep learning techniques*, *J. Ambient Intell. Humaniz. Comput.*, no. 0123456789, (2021)
- [5] M. Nakamura, J. Wang, S. Phea, and A. Ben Abdallah, *Comprehensive Study of Coronavirus Disease 2019 (COVID-19) Classification based on Deep Convolution Neural Networks*, *SHS Web Conf.*, **102**, 1–5, (2021)
- [6] M. E. H. Chowdhury *et al.*, *Can AI Help in Screening Viral and COVID-19 Pneumonia?*, *IEEE Access*, **8**, 132665–132676, (2020)
- [7] S. Sarker, L. Tan, W. Ma, S. Rong, O. Benjamin Kwapong, and O. Famous Darteh, *Multi-Classification Network for Identifying COVID-19 Cases Using Deep Convolutional Neural Networks*, *J. Internet Things*, **3**, no. 2, 39–51, (2021)
- [8] R. Yamashita, M. Nishio, R. Do, and K. Togashi, *Convolutional Neural Networks: An Overview and Its Applications in Pattern Recognition*, *Smart Innov. Syst. Technol.*, **195**, 21–30, (2021)
- [9] C. Dias *et al.*, *Using the Choquet Integral in the Pooling Layer in Deep Learning Network*, *Springer Int. Publ. AG*, **831**, 144–154, (2018)
- [10] M. Siddhartha and A. Santra, *COVIDLite: A depth-wise separable deep neural network with white balance and CLAHE for detection of COVID-19*, *J. Comput. Methods Programs Biomed.*, 1–25, (2020)
- [11] J. Lee, S. R. Pant, and H.-S. Lee, *An Adaptive Histogram Equalization Based Local Technique for Contrast Preserving Image Enhancement*, *Int. J. Fuzzy Log. Intell. Syst.*, **15**, 1, 35–44, (2015)