Custom convolutional neural network with data augmentation to predict Pneumonia COVID19

1st Budi Dwi Satoto
Information System Department
Engineering Faculty
Trunojoyo University
Madura, Indonesia
budids@trunojoyo.ac.id
https://orcid.org/0000-0002-1919-0540

2nd Mohammad Imam Utoyo
Applied Mathematics Department
Sains and Technology Faculty
Airlangga University
Surabaya, Indonesia
m.i.utoyo@fst.unair.ac.id
https://orcid.org/0000-0003-2292-8443

3rd Riries Rulaningtyas
Physics Department
Sains and Technology Faculty
Airlangga University
Surabaya, Indonesia
riries-r@fst.unair.ac.id

4th Eko Budi Koendhori
Microbiology department
Medical Faculty
Airlangga University
Surabaya, Indonesia
dr_eko@fk.unair.ac.id

2020 3rd International Conference on Biomedical Engineering (IBIOMED) | 978-1-7281-7156-2/20/$31.00 ©2020 IEEE | DOI: 10.1109/IBIOMED50285.2020.9487567

***Abstract*—Coronavirus is named because of the structure of the crown on its body surface. The effect of the Coronavirus for sufferers is a disturbance in the respiratory system. Several days later, the disruption from the lung infection got worse. To identify the cause of this disease, the doctor performs a computed tomography scan and manually observes the changes that occur in the lungs through an X-ray. Image identification using machine learning is the latest trend these days to assist medical analysts. If the number of patients treated is large enough, this is very helpful in the analysis. The choice of Convolutional Neural Network is due to the many architectural algorithms being developed at this time. This method works with multiple layers. But the drawback is that the computation time for the training process takes a long time. The purposed way in this research is a custom layer using 18-34 layers. There is four class in the test, namely Normal lung conditions, COVID19, bacterial pneumonia, and viral pneumonia. Data augmentation is used to add variation to data. The results showed that the method offered could be used to identify pneumonia with an average identification accuracy of 98.7% - 100%. The average value of error the MSE 18-34 layer is 0.0539, RMSE 0.1981, and MAE 0.0319. The average consumption time for the training process is 2.25 seconds. The best accuracy calculation is obtained at 34 layers with the Adaptive Moment Estimation optimizer with a computation time of around 1 minute 48 seconds.**

Keywords— COVID19, Pneumonia, Custom convolutional Neural Network, Data Augmentation

I. INTRODUCTION

The city of Wuhan China, at the end of 2019, became the first case of the discovery of the Coronavirus. More than 200 countries in the world have reported cases of exposure to the Coronavirus. A droplet from a person who is positive for COVID-19 and to a healthy human. Viruses cannot survive and reproduce on their own because they are looking for a host to live. They stay on other living organisms for energy that can be used for replication [1].

In mild infections, the sufferer only experiences sore throat, nasal congestion, and low-grade fever. Patients who experience fever and cough are moderate symptoms. The Coronavirus replicates to travel in the respiratory tract and enter the bronchial tubes. This condition can cause inflammation, which will result in a dry cough. In severe pulmonary infections, this virus leaves the bronchial tubes and enters deep into the lungs. As a result, the tissues involved in oxygen exchange are weak [2].

A series of X-Ray scans begin with 44-year-old Chinese patients. He died of the corona shows how the virus is damaging the lungs.The abnormality that doctors identified in the scan of corona patients was similar to that found in patients. From the results of a computer tomography scan, healthy lungs will be dark. White dots that appear are nodules or small growths in the lung tissue. Patients with COVID-19 infection will appear in white spots or patches on the lungs. The X-ray image performed looks like a crack in shattered glass. What is visible is fluid in the lung space. The only way to tell the difference is to pay attention to their shape and distribution [3].

X-ray research conducted by Sohaib Asif 2020 used an X-Ray image dataset consisting of 3 classes, namely COVID19, pneumonia virus, and normal conditions. The CNN method used is the Inception V3 model with transfer learning. Pre-processing is done by resizing the image to fit with the input layer. The classification accuracy results show the training accuracy is more than 98%, and the validation accuracy is 93%. In other research, Tulin Ozturk in 2020 conducted research using The DarkNet model for 2-3 class of pneumonia COVID19. The analysis applies different filters for each layer with 17 convolutional layers. In the classification process for class binaries, the resulting accuracy is 98.08%, while in multi-class, the accuracy is 87.02% [4].

Furthermore, X-ray analysis research was carried out by Makris (2020). The X-ray images used are mixed images available to the public with three output classes. Performance comparisons are made on several architectures. The results showed that the VGG16 and VGG19 designs were the best with an accuracy of 95%. Efficiency is being obtained using the NasnetLarge Model, with an average value of 81%. MobileNetV2 and densenet cannot exceed accuracy with 40% and 38% accuracy, respectively [5]. In his research (2020), Ioannis also uses a set of X-ray image data. There are three

classes and two data scenarios. VGG19 and MobileNet v2 used to compile the data. Suppose the False negative (FN) is low, then optimal results for accuracy can be obtained. The result is that significant biomarkers related to Covid-19 can be done using X-ray imaging processing. The result is that the best value obtained is 96.46% specificity, 98.66% sensitivity, and 96.78% accuracy [6].

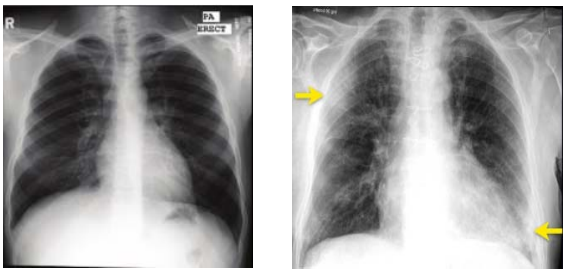
This proposed research analyzes X-rays of COVID19 pneumonia patients also using convolutional neural networks. GAP, with previous research, is that previous researchers still use the existing CNN architecture, and most research do not take into account the time of the training process. CNN has high accuracy, but the computation time takes a long time because the learning is done in-depth. The novelty offered is the addition of auto contacts and the use of custom layers with data augmentation. Auto contrast aims to clarify objects, and the Custom layer aims to reduce computation time. The use of development is to increase data variation because X-rays of COVID19 patients are not widely available on the internet. This research focuses on the lung condition of pneumonia sufferers with four leading causes, namely COVID19, normal lung conditions, lung conditions due to bacterial infection, and lung conditions due to viruses. Some of the future research that can be done include the use of K-Fold validation, optimization to get the typical parameter values, and parameter calculations values.

II. Literature Review and Methods

The literature review contains definitions and methods to be used

A. COVID-19 Virus

The size of this virus has a diameter of only 20 nanometers, which is lower than ribosomes. Based on direct observation using an electron microscope, with ultra-thin sections of infected living things [7]. Viruses are transmitted through a series of layers of different sizes. Meanwhile, the tail functions to make contact with the body of the organism it attacks [8].



(a) (b)

Fig. 1. Lung Condition (a) Normal (b) COVID19

The difference is shown in Fig 1. White patches are more visible in patients with the identification of COVID19.

B. Convolutional Neural Network

Convolutional Neural Neural works by arranging several variables into a target so that the objectives of the technique can be achieved [9]. The architecture of CNN is divided into two major parts, Feature Learning and Classification.

The underlying architecture is shown in Fig 2 [10].

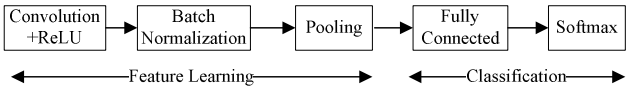


Fig. 2. The architecture of Convolutional neural network

The types of architecture include AlexNet, Googlenet, Resnet, VGG, Densenet, Squeezenet [11]. The difference lies in the number of layers used. More Layers will increase the number of parameters. It's just that it will add computing time [12]. The convolutional filter will shift by one of the stride values is 1. The stride is a determinant of the filter shift of N pixels in the filter convolution process. The transformation is done horizontally and vertically [13]—the layer relationship in performance expressed in Equation (1).

$$Number\ of\ neurons\ \text{fit} = \frac{(W-F+2ZP)}{S+1} \quad (1)$$

Where S is the stride of its application, F as the size of the filter, W is the width and height of the input image, and the number of zero paddings (ZP) [14]. The pooling layer is a layer that reduces the dimensions of the feature map—the formula Pooling layer expressed in Equation (2).

$$P_{out} = \frac{n_h-f+1}{s} x \frac{n_w-f+1}{s} x n_c \quad (2)$$

With P_{out}= pooling output, n_h = height of feature map, n_w = width of the feature map, n_c = number of channels, f = filter size, s = stride [15]. The pooling layer speeding up the computation because the parameters that have to be updated are fewer and overcome overfitting.[16].

C. Data augmentation

Data augmentation is the process of modifying images so that the computer will detect that the modified model is a different image. But humans can still know that the modified image is the same picture [17]. Shears expressed in Equation (3).

$$A = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \quad (3)$$

s defines the amount that I is sheared, and it is in the range of [0.1, 0.35]. Augmentation can improve the CNN model accuracy that trained [18]. Development performed in this research is to reverse the image horizontally, zoom in randomly, with a maximum zoom of 50% of the image size. Rotate pictures randomly with a maximum degree of 90° [19].

D. Accuracy

A confusion matrix is a matrix that is arranged based on the actual prediction class error [20]. Models that have undergone the training process will issue predictions. The Root Mean Square (RMSE) expressed in Equation (4) [21].

$$RMSE = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(Z_j - \hat{Z}_j)^2} \quad (4)$$

With N=number of the sample image, Z_j= actual label image, $\hat{Z}_j$ = predicted label image. Another calculation uses Mean absolute error (MAE) [22]. It expressed in Equation (5).

$$MAE = \frac{1}{n}\sum_{j=1}^{n}|y_j - \hat{y}_j| \quad (5)$$

MAE measures the magnitude of the average error in a series of predictions against the actual, without considering the direction. [23]

III. RESEARCH METHODS

The research method used is as illustrated in the flowchart below:

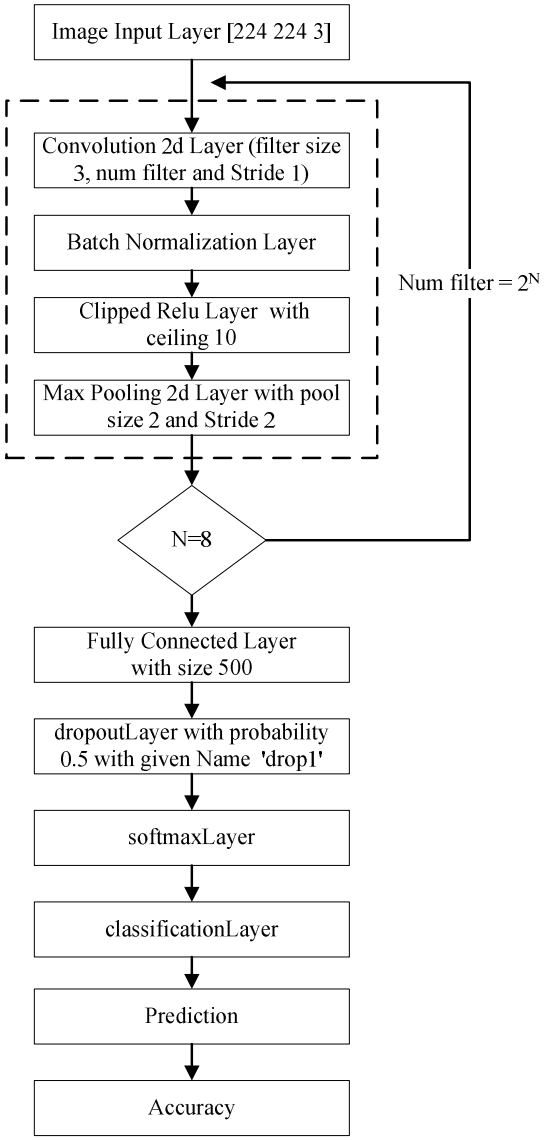


Fig. 3. Customize layer of Convolutional neural network

As shown in Fig 3, this research begins by preparing preliminary data in the form of X-Ray data obtained from the internet. Before entering CNN, the data is processed first using automatic contrast and resizing to format [224 224]. The convolutional layer process followed by the filter continues at the collection layer, which serves to get a feature map. A multidimensional array is a feature map that results from a feature extraction layer - reconfiguring the feature map into a vector. The software used is the MATLAB 2019a. The hardware used is a Laptop Core i-7 with 8GB of RAM with an NVIDIA GTX1050 Graphics Processing Unit (GPU).

Secondary data is available on the internet that can be used in this research. The Lung Condition dataset address used is https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia. Total File Size of 1GB. Picture of COVID X-Ray was taken from https://github.com/ieee8023/covid-chestxray-dataset/tree/master/images. The specifications are patients who are identified as having eight classes of pulmonary pain. In this research, training does use four types. The X-Ray data image files contain 43 COVID19, 234 Normal, 242 Pneumonia bacterial, and 148 Pneumonia virus [24].

IV. RESULT AND DISCUSSION

This stage contains the steps of Pre-processing, CNN, Pre-trained and data augmentation, and accuracy.

A. Initial process

The necessary preparation is to adjust the image dimensions similar to the aspects of the Convolutional neural network. This research uses 224x224 dimensions with one channel of the color. Dataset has various dimensions of size because data sources might be taken from different X-Ray photo machines. Data Image is shown in Fig 4.

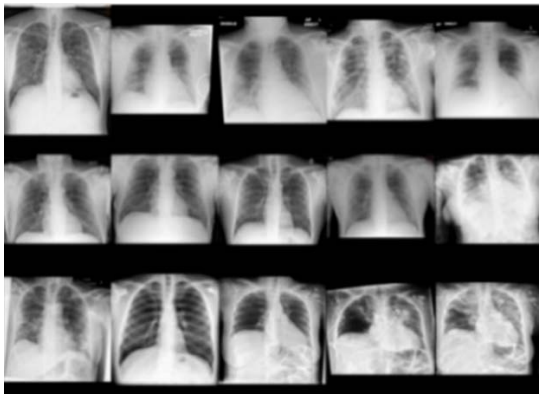
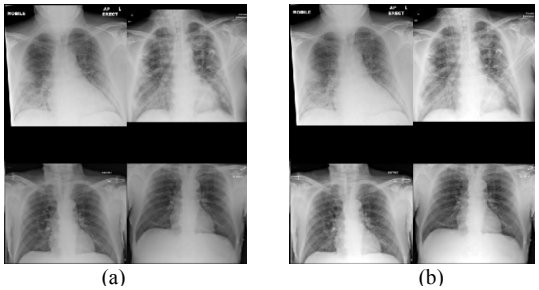


Fig. 4. Image input file

The next step is to run auto contrast to increase the intensity of the differentiation between the object and the background. The formula used is shown in Equation (6).

$$f_{ac}(a) = (a - a_{low}).\frac{255}{a_{high} - a_{low}} \quad (6)$$

with $f_{ac}(a)$= the mapping function. The gray-scale image has the full range from 0–255 using eight depth of bits. Output auto contrast has shown in Fig 5.

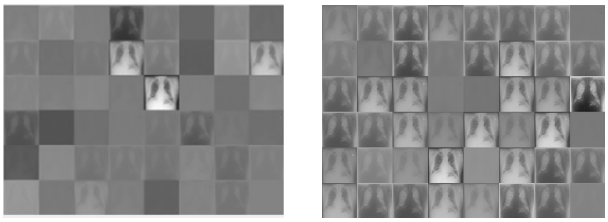


(a) (b)

Fig. 5. Pre-processing (a) before autocontrast (b) after autocontrast

B. Convolutional Layer and Feature Maps

The feature maps using padding "same" means the assumption that the output feature map size is the same as the input feature map. The goal is to get a planned feature with the same size as the input layer. Visualization of the Batch normalization layer shown in Fig 6.



(a) (b)

Fig. 6. Layer Convolution and Batch Normalization

C. Data augmentation

Image datastore works by enlarging training data sets, validation data, test data, and prediction data. The operation of resizing, rotating, and reflecting is carried out at this stage. The image data augmentation result shows in Fig 7. Scaling do by Equation (7).

$$A = \begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix} \quad (7)$$

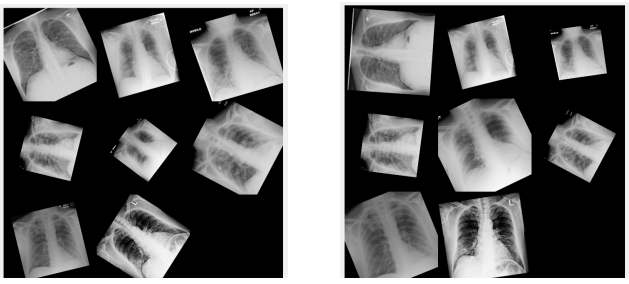


Fig. 7. Augmentation: Scaling and Rotation

Another technique uses for data augmentation is rotation. Each image I is rotated, represented by the following affine transformation. It expressed in Equation (8).

$$A = \begin{pmatrix} \cos\theta & -sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad (8)$$

While θ is the degree, the value is between 10 and 175 degrees. In this research used random rotation options of 90, arbitrary Scale 0.5, Random X Shear [-45 45], and Random Y Shear [-60 60]. The output is shown in Fig 8.

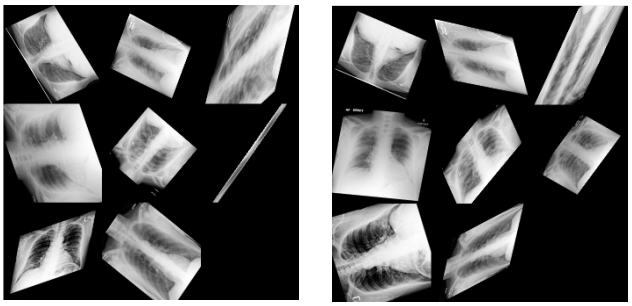


Fig. 8. Augmentation: Shear and reflection

D. Training

The training stage, a scenario using 18-34 layer. Options set Initial learning rate at $3x10^{-4}$, max epoch 30, iteration 100, validation Frequency 30 and minibatch 32. Training optimization used is rmsprop (root mean square propagation), sgdm (stochastic gradient descent with momentum), and adam (adaptive moment estimation). The training process using optimization shown in Fig 9.

E. Accuracy and Error classification

Confusion matrix can be used to determine classification accuracy by comparing testLabels and predicted Labels of image validation. F1-score expressed in Equation (9).

$$F1\ score = \frac{2x(RecallxPrecision)}{Recall+Precission} \quad (9)$$

Confusion Matrix F1 Score is a comparison of weighted average precision and recall written in Eq. (9). Table 1 shows the classification error calculation using Eq. (4) and Eq (5). The results show that at layer 34, the smallest MSE, RMSE, and MAE values are obtained. It is influenced by the stability of the training process because there are more convolutional layers. The average error cost of the MSE 18-34 layer is 0.0539, RMSE 0.1981, and MAE 0.0319. The average consumption time for the training process is 2.25 seconds.

TABLE I. RESULT WITH TRAINING OPTIMIZATION

Custom Layer	Output Class	Optimizer	Training Accuracy	Time consuming (minute)	Precision	Recall	F-1 Score	MSE	RMSE	MAE
18 layer	4	Adam	99,70%	01:33	0.9979	0.9979	0.9979	0,0085	0,0925	0,0043
		SGDM	99,40%	01:19	0,9939	0,994	0,9939	0,0162	0,1274	0,0095
		rmsprop	98,35%	02:07	0,9845	0,9849	0,9848	0,0283	0,1682	0,0198
22 layer	4	Adam	99,55%	03:05	0,9758	0,9758	0,9769	0,0755	0,2748	0,0413
		SGDM	99,25%	03:24	0,9968	0,9968	0,9968	0,0096	0,0979	0,0053
		rmsprop	96,85%	02:45	0,9758	0,9769	0,9765	0,0755	0,2748	0,0413
26 layer	4	Adam	97,90%	03:45	0,9826	0,9828	0,9828	0,0481	0,2193	0,0276
		SGDM	97,45%	01:53	0,975	0,9754	0,9755	0,0562	0,2370	0,0354
		rmsprop	95,05%	01:52	0,9538	0,9569	0,9549	0,1053	0,3245	0,0659
30 layer	4	Adam	99,55%	02:02	0,9962	0,9962	0,9962	0,0121	0,1098	0,0065
		SGDM	96,85%	01:54	0,9746	0,9746	0,9753	0,0758	0,2754	0,0422
		rmsprop	94,90%	01:55	0,9525	0,9533	0,9542	0,1229	0,3506	0,0715
34 layer	4	Adam	100,00%	01:48	1,0000	1,0000	1,0000	-	-	-
		SGDM	93,40%	01:11	0,9234	0,927	0,9245	0,1752	0,4186	0,1072
		rmsprop	100,00%	01:09	1,0000	1,0000	1,0000	-	-	-

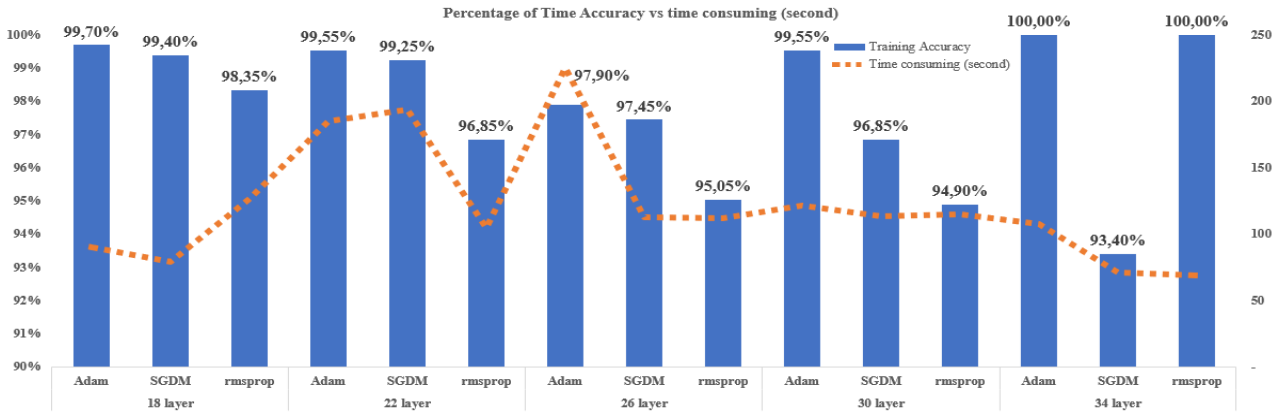


Fig 9. Comparison of Accuracy vs time consuming

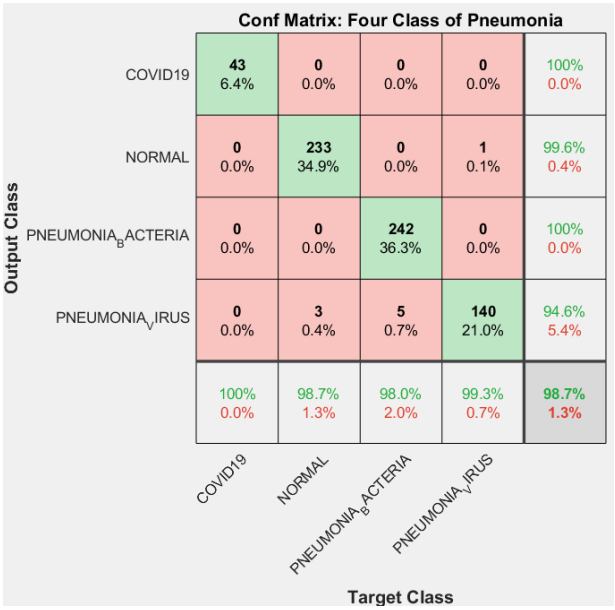


Fig. 10. Confusion matrix of pneumonia

From Fig 10, the Confusion matrix can be explained that from 40 COVID19 X-ray images, 43 can be classified correctly. Pneumonia bacteria, amounting to 242 files, five files are misclassified into the pneumonia virus class so that the accuracy is 98%. In the pneumonia virus, there is one file classify as Normal condition, so the efficiency is 99.6%. And from the total classification results obtained an average accuracy of 98.7%.

F. Prediction

Prediction is the stage that is used to test the data. This stage will match the image in the test data to the model formed from the results of the training process. The trials in Fig 11, show that the system can recognize pictures quite well. The first prediction results show the model can realize visions as pneumonia bacteria with an accuracy of 87.7% and 98.8%. The image is recognized as COVID19 98.2%. The healthy lung is 99.7%. In addition to the resulting efficiency, another thing to be concerned about is computational time. It is supported by the use of auto contrast to increase the difference between objects and the background so that the training process runs more stable.

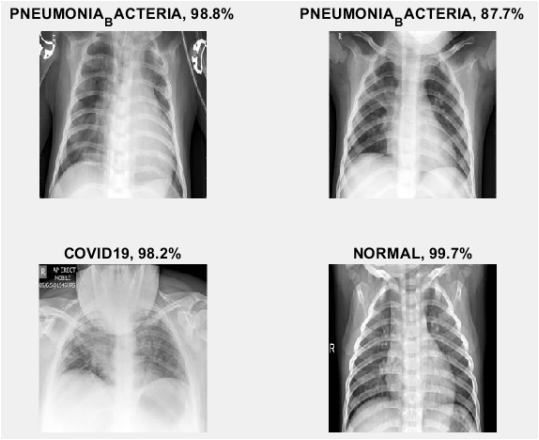


Fig. 11. Confusion matrix of pneumonia

G. Comparison of Research, Accuracy and time consuming

A comparison of the research that has been done by previous researchers using X-Ray images is shown in Table 2. In previous research, the existing architecture of CNN was still used. The number of output classes used is 3-4 classes. There is one researcher who uses pre-processing to improve the image. Image correction is also needed because some photos are sometimes taken with different devices so that they have different dimensions and the number of color channels.

TABLE 2. COMPARISON BETWEEN RESEARCH

No	Author	Pre process	Arch	No of layer	Aug	Avg Acc	Train time (min)
1	Makris [5]	#N/A	VGG	41-47	Avai lable	95%	33,33
2	Asif [4]	resize	Incep-tion V3	48	#N/A	98%	#N/A
3	Ozturk [25]	#N/A	Darknet	75	#N/A	98.08 %	#N/A
4	Ioannis [6]	#N/A	Mobile Net v2	53	#N/A	96.78 %	#N/A
5	Pro-posed Method	resize & Auto-contrast	Custom Layer	26-34	Avai lable	98.7 %	2,5

The comparison of layer performance shown in Fig 12 shows that using 34 layers with Adam's optimizer produces maximum accuracy. The best computational time for the training process needed is around 1 minute, 49 seconds. If the number of layers used is too small, the training process to find

accuracy performance runs too fast but tends to be unstable, training time is fast, but the resulting accuracy cannot be maximized. In comparison with existing architecture, the purposed method produces maximum efficiency with better computing time compared with alexnet and squeezenet. The best time needed is around 2 minutes and 30 seconds.

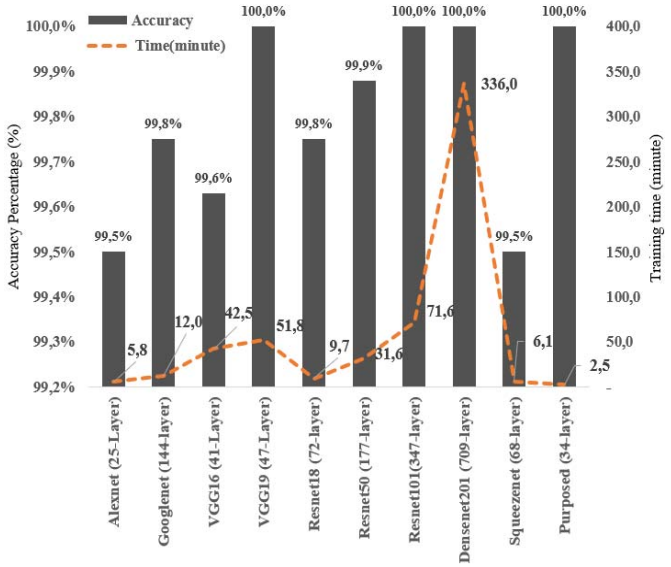


Fig. 12. Comparison of Architecture

V. CONCLUSION

The comparison results with other architectures show that the proposed custom layer can be used as a reference in the identification process. This research conducted observations on four output classes, namely bacterial pneumonia, Normal conditions, pneumonia COVID19, and viral pneumonia. The gap with previous research was in pre-processing, Number of Layer and time consumption. Proposals offered with custom layers 18-34 layers assisted with data augmentation can produce an average accuracy of 98.7-100% with computation time training processes ranging from one minute and 48 seconds. The average value of the MSE 18-34 layer is 0.0539, RMSE 0.1981, and MAE 0.0319, with average consumption time for the training process is 2.25 seconds. Research features that can be developed from this research are analysis with K-Fold validation and usings optimization to get the real parameter value.

ACKNOWLEDGMENT

Author Thanks to Dr. Soetomo Hospital in Surabaya for supporting primary data and the Ministry of Research, Technology and Higher Education, Grant 2019/2020, with contract number 812/UN3.14/PT/2020.

REFERENCES

[1] Ather A., B. Patel, N.B. Ruparel, "Coronavirus Disease 19 (COVID-19): Implications for Clinical Dental Care," Journal of Endodontics. Vol. 46, pp 584-595, 2020.

[2] Chen X., W. Wei, J. Wang, "Tide drove microbial dynamics through virus-host interactions in the estuarine ecosystem," Water Research. Vol. 160, pp 118-129, 2019.

[3] Ma L., X. Liu, Y. Gao, "A new method of content-based medical image retrieval and its applications to CT imaging sign retrieval," Journal of Biomedical Informatics. Vol. 66, pp 148-158, 2017.

[4] Asif S., Y. Wenhui, H. Jin, "Classification of COVID-19 from Chest X-ray images using Deep Convolutional Neural Networks," medRxiv. Vol. pp 2020.05.01.20088211, 2020.

[5] Makris A., I. Kontopoulos, and K. Tserpes, "COVID-19 detection from chest X-Ray images using Deep Learning and Convolutional Neural Networks," Vol. pp, 2020.

[6] Apostolopoulos I.D. and T.A. Mpesiana, "Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," Physical and Engineering Sciences in Medicine. Vol. 43, pp 635-640, 2020.

[7] Li X., M. Geng, Y. Peng, "Molecular immune pathogenesis and diagnosis of COVID-19," Journal of Pharmaceutical Analysis. Vol. 10, pp 102-108, 2020.

[8] Grove J. and M. Marsh, "The cell biology of receptor-mediated virus entry," The Journal of cell biology. Vol. 195, pp 1071-1082, 2011.

[9] Perera P. and V.M. Patel, "Learning Deep Features for One-Class Classification," IEEE Transactions on Image Processing. Vol. 28, pp 5450-5463, 2019.

[10] Saikia AR, K. Bora, L.B. Mahanta, "Comparative assessment of CNN architectures for classification of breast FNAC images," Tissue and Cell. Vol. 57, pp 8-14, 2019.

[11] Sandoval C., E. Pirogova, and M. Lech, "Two-Stage Deep Learning Approach to the Classification of Fine-Art Paintings," IEEE Access. Vol. 7, pp 41770-41781, 2019.

[12] Yan Z., Y. Zhan, S. Zhang. "Chapter 4 - Multi-Instance Multi-Stage Deep Learning for Medical Image Recognition", in Deep Learning for Medical Image AnalysisAcademic Press, 2017: p. 83-104.

[13] Yepez J. and S. Ko, "Stride 2 1-D, 2-D, and 3-D Winograd for Convolutional Neural Networks," IEEE Transactions on Very Large Scale Integration (VLSI) Systems. Vol. 28, pp 853-863, 2020.

[14] Risum A.B. and R. Bro, "Using deep learning to evaluate peaks in chromatographic data," Talanta. Vol. 204, pp 255-260, 2019.

[15] Yang W., Q. Liu, S. Wang, "Down image recognition based on deep convolutional neural network," Information Processing in Agriculture. Vol. 5, pp 246-252, 2018.

[16] Kamnitsas K., C. Ledig, V.F.J. Newcombe, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," Medical Image Analysis. Vol. 36, pp 61-78, 2017.

[17] Al-Qerem A., "An efficient machine-learning model based on data augmentation for pain intensity recognition," Egyptian Informatics Journal. Vol. pp, 2020.

[18] Mikołajczyk A. and M. Grochowski, "Data augmentation for improving deep learning in image classification problem." in 2018 International Interdisciplinary Ph.D. Workshop (IIPhDW), Vol. 117-122.2018.

[19] Bayr U. and O. Puschmann, "Automatic detection of woody vegetation in repeat landscape photographs using a convolutional neural network," Ecological Informatics. Vol. 50, pp 220-233, 2019.

[20] Kaur A. and I. Kaur, "An empirical evaluation of classification algorithms for fault prediction in open source projects," Journal of King Saud University - Computer and Information Sciences. Vol. 30, pp 2-17, 2018.

[21] Popoola S.I., A. Jefia, A.A. Atayero, "Determination of Neural Network Parameters for Path Loss Prediction in Very High-Frequency Wireless Channel," IEEE Access. Vol. 7, pp 150462-150483, 2019.

[22] Rémy N.M., T.T. Martial, and T.D. Clémentin, "The prediction of good physicians for prospective diagnosis using data mining," Informatics in Medicine Unlocked. Vol. 12, pp 120-127, 2018.

[23] Sharma P., Y.P.S. Berwal, and W. Ghai, "Performance analysis of deep learning CNN models for disease detection in plants using image segmentation," Information Processing in Agriculture. Vol. pp, 2019.

[24] Toğaçar M., B. Ergen, and Z. Cömert, "COVID-19 detection using deep learning models to exploit Social Mimic Optimization and structured chest X-ray images using fuzzy color and stacking approaches," Computers in Biology and Medicine. Vol. pp 103805, 2020.

[25] Ozturk T., M. Talo, EA Yildirim, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," Computers in Biology and Medicine. Vol. pp 103792, 2020.