Covid-19 classification using X-Ray imaging with ensemble learning

T Siswantining *, and R Parlindungan

Department of Mathematics, Faculty of Mathematics and Natural Sciences (FMIPA), Universitas Indonesia, Depok 16424, Indonesia

*titin@sci.ui.ac.id

Abstract. Coronavirus (Covid-19) first appeared in Wuhan, December 2019, and continues to spread rapidly to other countries. One of the countries infected with the Covid-19 virus is Indonesia. In Indonesia, the spread of this virus is very fast. Therefore, we need a detection system to detect people who are infected with this virus or not. Rapid detection of Covid-19 can contribute to control the spread of this disease. Chest x-ray images are one of the first imaging techniques to play an important role in the diagnosis of Covid-19. This research data uses chest x-ray images dataset in the Covid-19 cases. The data used in this study were 170 images data with 130 data for training data and 40 for testing data. In this study, the Neural Network, Support Vector Machine (SVM), and Convolutional Neural Network (CNN) methods were used, then applied to Stacking which is one of the methods of Ensemble Learning. The results of this study indicate that the best accuracy is obtained from the Stacking model with an accuracy of 95%.

1. Introduction

Coronavirus (Covid-19) is a disease that occurs because of virus that first occurred in the city of Wuhan, China in December 2019 and continues to spread in various countries [1]. The virus that causes Covid-19 is known as the acute respiratory syndrome coronavirus 2 or SARS-CoV-2. Coronavirus is a large family of viruses that cause airborne disease such as Middle East Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS) [2]. They are experiencing a fever, cough, and breathing difficulties. About 30% of people affected by Middle East respiratory syndrome coronavirus has died [3]. The same as COVID-19 virus causes milder symptoms in about 82% of cases, while others have severe or critical symptoms. The number of coronavirus cases in the world to date is around 4,250,862 and 286,986 of them died and 1,525,193 of them recovered, while the number of infected patients to date is 2,438,683. While 98% of patients infected with this disease are in mild condition, while the remaining 2% are in serious or critical condition (source: https://www.worldometers.info/coronavirus/).

Indonesia is one of the countries infected with the COVID-19 virus. In fact, the spread of this virus occurs very quickly. Based on data from the Task Force for the Acceleration of Handling COVID-19 of the Republic of Indonesia to date, the number of people infected with COVID-19 is 14,265, 2881 patients recovered, and 991 patients died. Signs of infection from this disease are respiratory problems, fever, cough and dyspnea. In some serious cases, infection from this virus can cause pneumonia, severe acute respiratory syndrome, septic shock, multiple organ failure, and death.

Rapid handling of Covid-19 can help control the spread of this disease. Detecting Covid-19 is usually associated with pneumonia symptoms and a chest x-ray image test. Chest x-ray images are one of the first imaging techniques to play an important role in the diagnosis of Covid-19. The results of the chest
x-ray are in the form of a dataset image that can be processed with the help of Machine Learning in determining whether the image is COVID-19 or not. This method can be used to facilitate early detection of people who are infected with the disease or not.

In this research, we used the Neural Network, Convolutional Neural Network and Support Vector Machine methods then ensemble using the stacking method to classify Coronavirus using chest x-ray images.

2. Method
This research stages are data in the form of chest x-rays, using the Neural Network, Convolutional Neural Network (CNN), and Support Vector Machine (SVM) methods which are then ensemble learning to classify Covid-19 disease or not.

2.1. Dataset
The data used in this research were 170 image data with 130 data for training data and 40 for testing data. The image size is reduced to 50 x 50 pixels so that computing time is not too long.

2.2. Pre-processing Image
This step is taken to improve the image quality to make it ready for further processing. Enhanced and enhanced image will aid in detection accuracy and improve overall image quality. Following are the steps used in the pre-processing stage:

2.2.1. Noise Removal
Medical images are damaged by various types of noise. It is very important to get the right images to facilitate accurate observations for a given application. One of the most common techniques for removing noise is the Median Filter. Median filter is a 'non-linear' filtering technique used to remove salt and Pepper noise from grayscale image [4].

2.2.2. Image Enhancement
Image enhancement essentially improves the interpretability and provides 'better' input for other automated image processing techniques. The main purpose of image enhancement is to modify image attributes to make it more suitable for a particular task and observer. During this process, one or more image attributes are modified [5].

2.3. Neural Network
Neural networks is a models that resembles the neurons in the human brain. Results of this operation will be a parameter of the activation function which will be the output of the neuron. The first architecture has 3 neurons on the Input Layer and 2 Output Layer nodes. Between Input and Output, there is 1 Hidden Layer with 4 neurons. Figure 1 shows the architecture of NN.

![Figure 1. Multilayer perceptron (MLP) or fully-connected layer architecture](https://medium.com/@nadhifasofia/1-convolutional-neural-network-convolutional-neural-network-merupakan-salah-satu-metode-machine-28189e17335b)
2.4. **Support Vector Machine**

Support Vector Machine (SVM) is a method which is usually good used for classification (such as Support Vector Classification) and regression (Support Vector Regression) [6]. In classification modelling, SVM has a more mature and clearer mathematical concept compared to other classification techniques. SVM can also solve classification and regression problems linear and nonlinear. Hyperplane is a function that can be used to separate between classes. In 2-D the function used for classifications between classes is called line whereas, the function used for classifying between classes in 3-D is called plane similarly, while the function used for classifications in higher dimensional classrooms is called a hyperplane.

![Figure 2. The illustration of SVM](https://example.com/svm Diagram)

The hyperplane found by SVM is illustrated as in Figure 2, its position is in the middle between two classes, meaning that the distance between the hyperplane and data objects is different from the adjacent (outermost) class which is marked with an empty and positive round. SVM is a conventional classification, but the result of svm accuracy are still good in several classifications, not only maximizing margins but also minimizing existing errors [8].

2.5. **Convolutional Neural Network (CNN)**

Convolutional Neural Network is a Multi Layer Perceptron (MLP) that is used to process two-dimensional data. CNN is a type of Deep Neural Network that is widely used in implementing image data. If analogous to using a human face, the first layer is a reflection of strokes in different directions, on the second layer features such as the shape of the eyes, nose, and mouth are visible, this is because the pooling is done from the first layer which is still a stroke- scratches, on the third layer a combination of the features of the eyes, nose, and mouth will be formed which will be inferred by the face of a certain person.

![Figure 3. Architecture of CNN](https://example.com/cnn Diagram)

Source: https://medium.com/@nadhifasofia/1-convolutional-neural-network-convolutional-neural-network-merupakan-salah-satu-metode-machine-28189e17335b
As with Neural Networks in general, CNN has several hidden layers from a single vector input. In Figure 3, the input is an image that is used as a single $32 \times 32$ vector. In each hidden layer, there are several neurons like the four C1 feature maps in the image. Neurons at C1 are linked to neurons at S1, and so on. The last layer connected to the previous hidden layers is called the output layer and represents the final class classification result. The overall scale in the object is very important so that the input does not lose its spatial information which will be extracted and classified by features. This will increase the accuracy and optimum level of the CNN algorithm. As in a cube which has a scale on length, width, and height. If you only use a regular Neural Network, it may only load the length and height scales. But CNN can contain all the information from the entire scale which can classify objects more accurately because it can use its wide scale too (which other two-dimensional Neural Networks may not see).

2.6. Ensemble Learning
The ensemble method is one of the algorithms in machine learning, where this algorithm looks for a solution with the best prediction than other algorithms, because this method uses several learning algorithms to get a better predictive solutions than several algorithms that can be obtained when using one only learning algorithm [9]. Predictive evaluations of ensembles usually require a lot of computation than single model predictive evaluations (single model), so this ensemble allows to compensate for poor learning algorithms by the performance of those computations. Fast algorithms such as decision trees are commonly used in these ensemble methods such as random forest, although slower algorithms can benefit from ensemble techniques as well.

2.7. Stacking
Stacking is a way to combine several models, with the concept of a meta learner. Unlike bagging and boosting, stacking allows combining models of different types. First-level learning is often generated by applying different learning algorithms. In the process of stacking training, a new data set must be created from the first stage classifier. Correct data at the first level classifier training stage, then the data used to generate new data sets used in the second level classifier training. This process carries a high risk of resulting in overfitting.

![Figure 4. Architecture of stacking](https://mc.ai/icu-survival-prediction-using-ensemble-learning-stacking/)

3. Result and Discussion
3.1. Neural Network
The training data that is used into the input layer is 130 data then goes to the first hidden layer of 300 neurons with the relu activation function. After that, it enters the second hidden layer of 200 neurons with relu activation function. The output from the second hidden layer to the output layer is 2 neurons because the data has 2 classes. A summary of the layers and parameters used in the NN model can be seen in Table 1.
Table 1. Summary of layers and parameters in the neural network model

| Layer (type) | Output Shape | Param # |
|--------------|--------------|---------|
| dense_15 (Dense) | (None, 300) | 2250000 |
| dense_16 (Dense) | (None, 200) | 60200 |
| dense_17 (Dense) | (None, 2) | 402 |

Total params: 2,310,902
Trainable params: 2,310,902
Non-trainable params: 0

Modeling is done with epoch = 160, batch size = 20 and validation split = 0.2.

Confusion Matrix and Statistics

| Actual | Predicted | 0 | 1 |
|--------|-----------|---|---|
| 0       | 88        | 0 | 0 |
| 1       | 39        | 0 | 0 |

Accuracy : 0.9769
95% CI : (0.9734, 0.9802)
No Information Rate : 0.5
P-Value [Acc > NIR] : <2e-16
kappa : 0.9462

Mcnemar's Test P-Value : 0.2482
Sensitivity : 0.9670
Specificity : 1.0000
Pos Pred Value : 1.0000
Neg Pred Value : 0.9286
Prevalence : 0.0000
Detection Rate : 0.9670
Detection Prevalence : 0.9670
Balanced Accuracy : 0.9835

'Positive' Class : 0

Figure 5. Evaluation of the Neural Network model from the training data

Confusion Matrix and Statistics

| Actual | Predicted | 0 | 1 |
|--------|-----------|---|---|
| 0       | 22        | 0 | 0 |
| 1       | 12        | 0 | 0 |

Accuracy : 0.925
95% CI : (0.7961, 0.9843)
No Information Rate : 0.5
P-Value [Acc > NIR] : 0.0001938
kappa : 0.8324

Mcnemar's Test P-Value : 1.0000000
Sensitivity : 0.9250
Specificity : 0.9211
Pos Pred Value : 0.9515
Neg Pred Value : 0.8571
Prevalence : 0.0000
Detection Rate : 0.9250
Detection Prevalence : 0.9250
Balanced Accuracy : 0.9245

'Positive' Class : 0

Figure 6. Evaluation of the neural network model from the testing data

Figure 5 and Figure 6 shows that the evaluation of the Neural Network model from the training data get an accuracy of 97.69% and from the testing data get an accuracy of 92.5%.
3.2. Convolutional Neural Network (CNN)

The training data used in the input layer on CNN is 130. For each activation function used on the CNN layer, it is relu. In the output layer there are 2 neurons because the data has 2 classes. A summary of the layers and parameters used in the CNN model can be seen in Table 2.

| Layer (type)           | Output Shape     | Param # |
|------------------------|------------------|---------|
| conv2d_29 (Conv2D)     | (None, 48, 48, 32) | 896     |
| max_pooling2D_6 (MaxPooling2D) | (None, 23, 23, 32) | 0       |
| dropout_32 (Dropout)   | (None, 23, 23, 32) | 0       |
| conv2d_28 (Conv2D)     | (None, 21, 21, 64)  | 18464   |
| conv2d_29 (Conv2D)     | (None, 19, 19, 64)  | 36992   |
| max_pooling2d_9 (MaxPooling2D) | (None, 0, 0, 64)    | 0       |
| dropout_34 (Dropout)   | (None, 0, 0, 64)    | 0       |
| flatten_4 (Flatten)    | (None, 1, 512)     | 0       |
| dense_18 (Dense)       | (None, 256)        | 1327360 |
| dropout_34 (Dropout)   | (None, 256)        | 0       |
| dense_19 (Dense)       | (None, 2)          | 512     |

Total params: 1,303,442
Trainable params: 1,303,442
Non-trainable params: 0

Modeling is done with epoch = 50, batch size = 20 and validation split = 0.2 to prevent overfitting.

Figure 7. Results of the evaluation of CNN model from the training data

Figure 8. Results of the evaluation of CNN model from the testing data
Figure 7 and Figure 8 shows that the evaluation of the CNN model from the training data get an accuracy of 96.92% and from the testing data get an accuracy of 87.5%.

3.3. Support Vector Machine (SVM)
In this paper, the kernel function used is the sigmoid function and the SVM type uses C-Classification.

![Figure 9. Evaluation of SVM model from the training data](image)

| Model                      | Accuracy (%) |
|----------------------------|--------------|
| Neural Network             | 92.5         |
| Convolutional Neural Network (CNN) | 87.5         |
| Support Vector Machine (SVM) | 87.5         |
| Stacking                   | 95           |

Figure 9 and Figure 10 shows that the evaluation of the SVM model from the training data get an accuracy of 89.23% and from the testing data get an accuracy of 87.5%.

3.4. Stacking
The stacking results of the three models with data testing can be seen in Table 3.
4. Conclusion
The ensemble learning method with stacking produces better accuracy than the three models with an accuracy of 95%. The accuracy results for the neural network method are 5% better than the accuracy results for the CNN method. This proves that the more layers used in modelling does not mean the result will be better. The accuracy results can be improved again if more data are processed to make models. The data processed is only 170 image data from a total of 5910 available image data. This is due to the limitations of the computer specifications used to make the model. The use of the results of this model still cannot be used in real life.

Acknowledgments
This study is supported by PUTI Proceedings Research Grant from University of Indonesia: NKB-962/UN.2RST/HKP.05.00/2020. The authors are thankful to all parties that involve in the process of writing this paper.

References
[1] Abbas A, Abdelsamea M M, and Gaber M M 2020 Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network arXiv preprint arXiv:2003.13815
[2] Narin A, Kaya C, and Pamuk Z 2020 Automatic detection of Coronavirus Disease (COVID-19) using X-ray images and deep convolutional neural networks arXiv preprint arXiv: 2003.10849, 1-17
[3] Bustamam A, Ulul E D, Hura H F A, and Siswantining T 2017 Implementation of hierarchical clustering using k-mer sparse matrix to analyze MERS-Cov genetic relationship AIP Conference Proceedings 1862 030142
[4] Patil R C, Bhalchandra A S 2012 Brain tumour extraction from MRI images Using Matlab. IJECSCE, ISSN: 2277-9477, Volume 2, issue 1
[5] Gadpayle P and Mahajani P S 2013 Detection and classification of brain tumor in MRI images International Journal of Emerging Trends in Electrical and Electronics
[6] Sarwinda D, Siswantining T, and Bustamam A 2018 Classification of diabetic retinopathy stages using histogram of oriented gradients and shallow learning International Conference on Computer, Control, Informatics and its Aplication (IC3INA) 83-87
[7] Hamidah, Rustam Z, Utama S, and Siswantining T 2020 Multiclass classification of acute lymphoblastic leukemia microarrays data using support vector machine algorithms Journal of Physics Conferences Series 1490
[8] Rustam Z, Syarifah M A, and Siswantining T 2019 Recursive particle swarm optimization (RPSO) schemed support vector machine (SVM) implementation for microarray data analysis on chronic kidney disease (CKD) IOP Conference Series: Materials Science and Engineering 546
[9] Bustamam A, Formalidin S, and Siswantining T 2018 Clustering and analyzing microarray data of lymphoma using singular value decomposition (SVD) and hybrid clustering AIP Conference Proceedings of The 3rd International Symposium on Current Progress in Mathematics and Science (ISCPMS 2016) 2023 020220