COVID-19 Chest X-Ray Classification Using Convolutional Neural Network Architectures

Devi Nurtiyasari
Department of Mathematics
Universitas Gadjah Mada
Yogyakarta, Indonesia
devi.nurtiyasari@mail.ugm.ac.id

Dedi Rosadi
Department of Mathematics
Universitas Gadjah Mada
Yogyakarta, Indonesia
dedirosadi@gadjahmada.edu

Corresponding Author: Dedi Rosadi

2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI) | 978-1-7281-8406-7/20/$31.00 ©2020 IEEE | DOI: 10.1109/ISRITI51436.2020.9315499

***Abstract*— World Health Organizations declared that Coronavirus Disease 2019 (COVID-19) outbreak pandemic in March 2020. Countries around the world are stepping up effort to halt the spread of this pandemic. Some countries are scrambling to tackle this virus by applying lockdown policy. As of 10 August 2020, there have been confirmed 19.718.030 total cases and 728.013 total deaths of COVID-19 [2]. COVID-19 detection is vital to decide the subsequent step in handling the patients. One strategy that may be applied for COVID-19 detection is classification approach primarily based totally on chest x-ray of the patients. Convolutional neural network has been successfully applied in practical applications. It is a type of machine learning which the model is designed to learn classification tasks directly from an image. It recognizes patterns directly from image pixel. These patterns are used to classify images and to eliminate the need of manual feature extraction. The classification provides outcomes with recall, precision, and accuracy had been respectively 94.99%, 95%, and 95.47% for model 1 and 97.73%, 95%, and 96.59% for model 2.**

Keywords*— *convolutional neural network, covid-19, neural network, pandemic, coronavirus disease,

I. Introduction

According to World Health Organization [1], Coronavirus Disease 2019 (COVID-19) is a disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). This virus is a family viruses that can cause common cold, SARS (Severe Acute Respiratory Syndrome), and MARS (Middle East respiratory syndrome). As it is a new strain of coronavirus, the antiviral is not yet found. World Health Organization declared the COVID-19 outbreak a pandemic in March 2020. Some countries are scrambling to tackle this virus by applying lockdown policy. As of 10 August 2020, there have been confirmed 19.718.030 total cases and 728.013 total deaths of COVID-19 [2].

The symptoms of COVID-19 are similar to the flu (influenza) or the common cold, i.e. fever, cough, shortness of breath, and breathing difficulties [1]. It can lead to pneumonia, multiple organ failure and even death in severe case. This is why testing is required to confirm if someone has COVID-19 or not. COVID-19 detection is vital to decide the subsequent step in handling the patients. One strategy that may be applied for COVID-19 detection is classification approach primarily based on chest x-ray of the patients. This image provides information on whether the lung is normal or not. These are interesting instances for medical image processing since innovations in deep learning and the increasing availability of massive annotated medical image datasets are leading to dramatic advances in automated knowledge of medical images [3].

Most lung image classification based on chest x-ray using neural network model with manual features extraction process. The preceding studies for lung image classification based on chest x-ray has been accomplished through the use of recurrent neural network for lung image and it provides accuracy result 81.33% [4]. This study then evolved through the use of wavelet for image denoising process, the model was called as wavelet recurrent neural network and it affords outcome with accuracy 84% [5]. Wavelet recurrent neural network used the wavelet to eliminate image noise in the preprocessing step and used the recurrent neural network for the classification step. These previouses research using manual features extraction process as the input for the classification process. Convolutional Neural Network (CNN) is special types of neural network. CNN is learning to perform classification tasks directly from an image and there is no need to do manual features extraction in this model. CNN can recognize patterns with extreme variability, and with robustness to distortions and simple geometric transformations [6]. The classification accuracy of Covid-19 Chest X-Ray using CNN is expected an increase as CNN model has been successfully applied in practical applications, such as the application of CNN for image classification that is provides accuracy 96.5% [7] and the application of CNN for image recognition that is provides accuracy 98.89% [8].

II. Convolutional Neural Network For Covid-19 Chest X-Ray Classification

In 1980, Fukushima designed Neocognition [9], which is a model of unsupervised artificial networks inspired by the visual biological system. LeCun, et al. [10] applied the model to solve handwritten digit recognition problems. Furthermore, LeCun et al. [11] introduced the term of *Convolutional Neural Network (CNN)*. CNN is a deep neural network in which neural systems keep up a various leveled structure by representing the feature learning and generalizing it into the common image problem. CNN isn't constrained to images, but to receives the advanced in natural language processing problems and speech recognition problems [12]. CNN is the advancement of Multi-Layer Perceptron (MLP) which is outlined to prepare grids data, one of which is utilized for image analysis and image processing. The feature learning of CNN is then used to distinguish between images from one another and to classify it.

Generally, CNN is not that different with neural networks. The neurons in CNN consist of weight, bias and activation functions. Convolutional layer in CNN comprises of neurons arranged to create a filter. CNN utilizes the convolution process by moving a certain size convolution kernel (filter) to an image that makes the computer then get new representative information from the multiplication of parts of the image with the filter used.

A. Architecture of CNN

CNN architecture is created and inspired by the way humans produce visual perception (visual cortex). The primary function of the visual cortex is to process visual information [13]. With this architecture, CNN can be trained to understand the details of an image, so that it can capture spatial and temporal dependencies in an image. CNN combines three architectural ideas to ensure some degrees of shift and distortions invariance: local receptive fields, shared weights (weight replication), and sometimes, spatial or temporal subsampling [10]. In general, the architecture of CNN is divided into 2 major parts, i.e. Feature Learning and Classification. The architecture of CNN can be seen in Fig. 1. This figure was adapted from the CNN architecture for handwritten recognition [11]. In this architecture, the input is in the form of image.

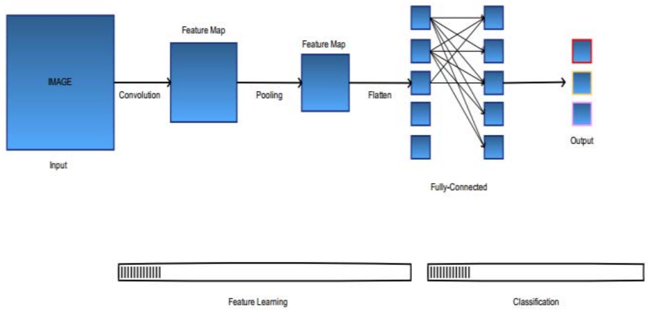

Fig. 1. CNN Architecture

Image is pixels framework which arranges in rows and columns (matrix). Pixels describe the location of rows and columns. It also has an intensity which is denoted as p(x,y). The image as the input on CNN will be read as matrix data by a computer, as shown in Fig. 2. If the CNN input is an RGB image, then the image will have 3 matrices, i.e. the matrix for the colors Red, Green, and Blue where the matrix value ranges between 0 up to 255 for each color in each pixel of the image. If the CNN input is a Graysclae image, then the image will have a matrix whose values range from 0 (black) to 255 (white). Meanwhile, if the CNN input is in the form of a black and white image, the image will have a matrix whose value has only 2 possibilities, i.e. 0 (black) and 1 (white). The point of any image processing technique is that it allows the machine to recognize the characteristics of the input image, including this CNN technique. In this case, the input image has not yet provide any information, so it is necessary to do the next step, i.e. Feature Learning and Classification.

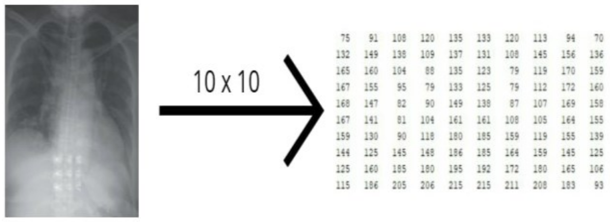

Fig. 2. Image Matrix

B. Feature Learning

There are several important points in the Feature Learning section, i.e. Convolution, Activation Functions and the Pooling Layer. The feature learning is rehashed a few times until an adequate feature map is obtained to continue the classification stage.

1) Convolution

At Convolutional layer, neurons arranged to create a filter that contains pixels. This filter will be moved all parts of the image. CNN utilizes the convolution process by moving a certain size convolution kernel (filter) to an image that makes the computer then get new representative information from the multiplication of parts of the image with the filter used. Technically, a dot product operation will be performed between the input and the value of the filter at each shift, so that it produces an output or so-called feature map. The preceding layer's (input layer) feature maps are convolved with learnable kernels and positioned via the activation function to form the output feature map [14]. The output map may combine convolutions with numerous input maps. Mathematically, it can be written as follows [14]:

$$x_j^l = \sum_{i \in M_j} x_i^{l-1} . x_{ij}^l + b_j^l \tag{1}$$

where M_j represents a selection of input maps, x_j^l is the output map j for convolution layer's l, and k is the kernel. The output map is spesified with an additive bias, but for a spesific output map, the input maps might be convolved with other recognizable kernels. The output x_j^l is activated using activation function f, then $x = f\left(x_j^l\right)$.

2) Activation Function

The activation function is at the stage between the convolution process and the pooling layer. At this stage, the convolution value is subject to the activation function. The activation function that is most often used in the CNN model is the rectified linear unit or ReLU [15] and is formulated as follows:

$$f(x) = \max(x, 0) \tag{2}$$

A ReLU of x is defined as the maximum value of 0 and x, which means that it will equals to 0 for any negative input value. Meanwhile, if the input value of the activation function is positive, then the output of the neuron is the input value of the activation itself.

3) Pooling Layer

The activation function is subsequently followed by additional convolution such as the pooling layer, fully connected layer and normalization layer. These refers as hidden layers because their input and output are covered by the activation function and final convolution. The pooling layer generally consist of a filter with spesific size and stride that shifts throughout the feature map area. Pooling commonly used is Max Pooling and Average Pooling. The reason of utilizing the pooling layer itself is to diminish the dimensions of the feature map (down-sampling), thereby speeding up the computation because fewer parameters need to be updated. In addition, the reason of utilizing pooling layer is to overcome over fitting.

C. Classification

The results of feature learning are feature maps in the form of multidimensional arrays, so in this case the flatten process must be carried out. Through the flatten process, a feature map reshape will be made from a multidimensional array into a vector so that it can be used as the fully-connected layer's input. Fully-connected layers are ordinary neural networks, i.e. Multi Layered Perceptron (MLP). At this layer, all activity neurons from the preceding layer are associated to neurons within the next layer. The fully connected layer points to process data so that it can be classified.

D. Diagnostic Test

Diagnostic test relates to the capacity of a test to separate between the target condition and wellbeing (health) [16]. The separative potential can be measured by the measures of accuracy, recall and precision. The probability that may show up in diagnostic test are shown in Table 1.

TABLE I
DIAGNOSTIC TEST

Measure	True Situation		**Total**
	Performance Indicator Present	Performance Indicator Absent	
Positive	True Positive (tp)	False Positive (fp)	tp+fp
Negative	False Negative (fn)	True Negative (tn)	fn+tn
Total	tp+fn	fp+tn	tp+fp+fn+tn

There are four outcomes for diagnostic tests, i.e. true positive (tp), false positive (fp), false negative (fn), and true negative (tn). The (tp) means that a sick patient is accurately recognized as sick, (fp) means that healthy patient is inaccurately recognized as sick, (fn) means that sick patient is inaccurately recognized as healthy, and (tn) means that healthy patient is accurately recognized as healthy. Recall refers to the percentage of relevant results that are correctly classified by the model. This is the relationship between (tp) and the total number of (tp) and (fn). Mathematically, it is calculated as follows:

$$recall = \frac{tp}{tp + fn} \tag{3}$$

Precision refers the percentage of the results which are relevant. This is the relationship between true positives and the total number of true positives and false positives. It is mathematically measured by:

$$precision = \frac{tp}{tp + fp} \tag{4}$$

The accuracy is the percentage of correct predictions result in the model among the total cases. It is a statistical measurement to define the wellness of the classficiation correctly identifies condition. The formula for quantifying accuracy is:

$$accuracy = \frac{tp + tn}{tp + tn + fp + fn} \tag{5}$$

III. RESULT AND DISCUSSION

The dataset used in this paper is open source. There are 97 positive Covid-19 chest x-ray dataset [17] and 121 negative/normal chest x-ray dataset [18]. These dataset divided as training data (174 images) and testing data (44 images). The CNN model will automatically train the x-ray image characteristic through the Feature Learning and Classification process. It will determine wether the x-ray image is positive Covid-19 chest x-ray or normal chest x-ray.

A. Model 1

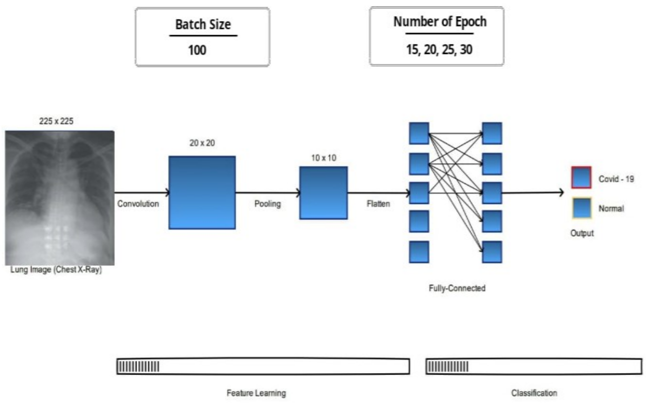

Fig. 3. Model 1

The fisrt CNN model (Fig. 3.) is using 225x225 image matrix, 20x20 convolution matrix , 10x10 pooling matrix with 100 batch size and the number of epoch is 15, 20, 25, and 30. The result of the process is shown in Table II.

TABLE II
DIAGNOSTIC TEST RESULT FOR MODEL 1

Number of Epoch	(tp)	(fp)	(fn)	(tn)
15	18	2	1	23
20	18	2	1	23
25	20	0	1	23
30	20	0	1	23

As shown in Table II, CNN model for Covid-19 classification based on chest x-ray provided result 23 true negatives and 1 false negative for all number epoch choosen. The learning process using 15 number of epoch gives the same result as the learning process using 20 number of epoch. Meanwhile, the learning process using 25 number of epoch gives the same result as the learning process using 30 number of epoch. The precision, recall and accurary of the model are calculated as shown in Table III. Based on Table III, the average recall, precision, and accuracy are respectively 94.99%, 95%, and 95.47%.

TABLE III
RECALL, PRECISION, AND ACCURACY OF MODEL 1

Number of Epoch	Recall	Precision	Accuracy
15	94.74%	90%	93.2%
20	94.74%	90%	93.2%
25	95.24%	100%	97.73%
30	95.24%	100%	97.73%

B. *Model 2*

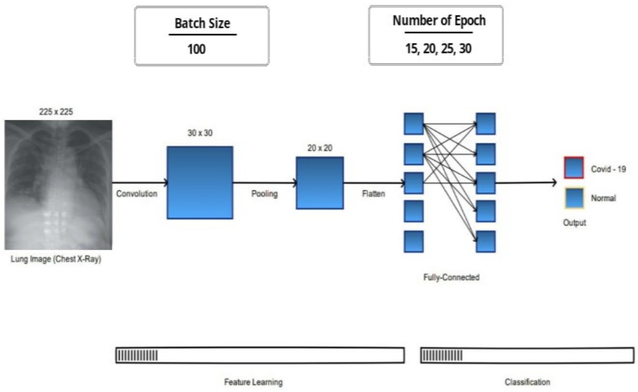

Fig. 4. Model 2

The second CNN model (Fig.4.) is using 225x225 image matrix, 30x30 convolution matrix, and 20x20 pooling matrix with 100 batch size and the number of epoch is 15, 20, 25, and 30. The result of the process is shown in Table IV.

TABLE IV
DIAGNOSTIC TEST RESULT FOR MODEL 2

Number of Epoch	(tp)	(fp)	(fn)	(tn)
15	18	2	0	24
20	19	1	0	24
25	20	0	2	22
30	19	1	0	24

As shown in Table IV, CNN model for Covid-19 classification based on chest x-ray provided the same result for learning using 20 number of epoch and 30 number of epoch. The precision, recall and accurary of the model are calculated and shown in Table V. Based on Table V, the average recall, precision, and accuracy are respectively 97.73%, 95%, and 96.59%.

TABLE V
RECALL, PRECISION, AND ACCURACY OF MODEL 2

Number of Epoch	Recall	Precision	Accuracy
15	100%	90%	95.45%
20	100%	90%	97.73%
25	90.91%	100%	95.45%
30	100%	95%	97.73%

IV. CONCLUSION

The procedures of Covid-19 chest X-Ray classification using Convolutional Neural Network was defining the input data, i.e. training data and testing data. The input data is used to determine the learning features and classification process in CNN architecture. The CNN model will automatically train the X-Ray image characteristic through the convolutional layer and the pooling layer. The classification process starts at the fully connected layer. This will determine whether the x-ray image is positive Covid-19 or normal.

The first model is using 225x225 image matrix, 20x20 convolution matrix , 10x10 pooling matrix with 100 batch size and the number of epoch is 15, 20, 15, and 30. It afforded outcomes with recall, precision, and accuracy were respectively 94.99%, 95%, and 95.47%. The second model is using 225x225 image matrix, 30x30 convolution matrix, and 20x20 pooling matrix with 100 batch size and the number of epoch is 15, 20, 25, and 30. It afforded outcomes with recall, precision, and accuracy were respectively 97.73%, 95%, and 96.59%.

This research used an open dataset and because of these limitations, this research only uses 97 positive chest x-rays and 121 normal chest X-rays. However, this model will provide better results and accuracy by using larger dataset, and need to be supported by clinical study.

REFERENCES

[1] World Health Organizations. Coronaviruse: Situation Report. (2020). [Online]. Available: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200423-sitrep-94-covid-19.pdf. Accessed: August 10, 2020

[2] World Health Organizations. Novel Coronavirus 2019. (2020). [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019. Accessed: August 10, 2020

[3] Summers R.M., “Deep Learning and Computer-Aided Diagnosis for Medical Image Processing: A Personal Perspective,” in Deep Learning and Convolutional Neural Networks for Medical Image Computing. New York, NY, USA: Springer, 2017, pp: 3-10.

[4] D. Nurtiyasari, “The Application of Recurrent Neural Network Model for Lung Cancer Nodule Classification,” unpublished.

[5] D. Nurtiyasari, D. Rosadi, and Abdurakhman, "The application of Wavelet Recurrent Neural Network for lung cancer classification," in *3rd International Conference on Science and Technology - Computer (ICST)*, Yogyakarta, Indonesia, 2017, pp. 127-130.

[6] LeCun,Y. LeNet-5, Convolutional Neural Network. (2015). [Online]. Available: http://yann.lecun.com/exdb/lenet/. Accessed: April 1, 2020

[7] M. Hussain, J. Bird, and D. Faira, “A Study on CNN Transfer Learning for Image Classification,” in Advances in Computational Intelligence System: 18th Annual UK Workshop on Computational Intelligence. New York, NY, USA: Springer, 2018, pp: 31-33.

[8] M.O. Weilong, L. Xiaoshu, Z.Yexiu, and J. Wenjie, “Image Recognition Using Convolutional Neural Network Combined with Ensemble Learning Algorithm,” in Journal of Physics: Conference Series, 2019. [Online]. Available: https://iopscience.iop.org/article/10.1088/1742-6596/1237/2/022026/pdf

[9] K. Fukushima, “Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position,” in Biological Cybernetics 36, 1980, pp. 193-202.

[10] Y. LeCun, B. Boser, J. Denker, D. Henderson, R. Howard, and L. Jackel, “Handwritten Digit Recognition with a Back-Propagation Network,” in Advances in Neural Information Processing System, vol.2, 1989, pp. 396-404. [Online]. Available: https://papers.nips.cc/paper/293-handwritten-digit-recognition-with-a-back-propagation-network

[11] Y. LeCun and J. Bengio, “Convolutional Network for Image, Speech, and Time Series,” in The Handbook of Brain Theory adn

Neural Networks, 1995, vol 3361, no 10. [Online]. Available: http://yann.lecun.com/exdb/publis/pdf/lecun-bengio-95a.pdf

[12] N.K. Manaswi, Deep Learning with Applications Using Python: Chatbots and Face, Object, and Speech Recognition With TensorFlow and Keras. New York, NY, USA: Apress, 2018, pp. 91-92.

[13] H. Trevor, M. Navid, and T. Prasnna, Neuroanatomy, Visual Cortex. Petersbug, Florida, USA: Statpearls, 2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK482504/

[14] S. Liu, C. Zhang, and J. Ma, "CNN-LSTM Neural Network Model for Quantitative Strategy Analysis in Stock Market," Neural Information Processing: The 24 International Conferences ICONIP, New York, NY, USA: Springer, 2017, pp: 198-206.

[15] S., Skansi, Introduction to Deep Learning: From Calculus to Artificial Intelligence. New York, NY, USA: Springer, 2018, pp: 123-124.

[16] C.M., Leutenegger, "Test Performance," in Interpretation of Equine Laboratory Diagnostic. Hoboken, New Jersey, USA: John Wiley & Sons, 2018, pp: 27-28.

[17] J.P. Cohen, P. Morrison, and L. Dao. Covid-19 Image Data Collection. (2020). [Online]. Available: https://github.com/ieee8023/covid-chestxray-dataset. Accesed: March 23, 2020

[18] P. Mooney. Chest X-Ray Images. (2020). [Online]. Available: https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge. Accesed: March 23, 2020