Available online at http://scik.org

Commun. Math. Biol. Neurosci. 2020, 2020:42

https://doi.org/10.28919/cmbn/4765

ISSN: 2052-2541

DETECTION OF COVID-19 CHEST X-RAY USING SUPPORT VECTOR MACHINE AND CONVOLUTIONAL NEURAL NETWORK

DIAN CANDRA RINI NOVITASARI[1,2,*], RIMULJO HENDRADI[2], REZZY EKO CARAKA[3,6,*], YUANITA RACHMAWATI[4], NURUL ZAINAL FANANI[5], ANANG SYARIFUDIN[1], TONI TOHARUDIN[6], RUNG CHING CHEN[3]

[1]Mathematics Department, Faculty of Science and Technology, UIN Sunan Ampel Surabaya, Indonesia, 60237

[2]Mathematics Department, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia, 60115

[3]Department of Information Management, College of Informatics, Chaoyang University of Technology, Taiwan (ROC), 41349

[4]Biology Department, Faculty of Science and Technology, UIN Sunan Ampel Surabaya, Indonesia, 60237

[5]Informatic Department, Politeknik Negeri Jember, Indonesia, 68101

[6]Department of Statistics, Padjadjaran University, West Java, Indonesia, 45363

Abstract This study aims to detect whether patients examined are healthy, Coronavirus positive, or just have pneumonia based on chest X-ray data using Convolutional Neural Network method as feature extraction and Support Vector Machine as a classification method or called Convolutional Support Vector Machine. Experiments carried out were comparing the kernel used, feature selection methods, architecture in feature extraction, and separated classes. Our instrument reached the accuracy of 97.33% in the separation of 3 classes (normal, pneumonia, COVID19) and 100% in the separation of 2 classes, that is (normal, COVID19) and (pneumonia, COVID19), respectively. Based on

*Corresponding authors

E-mail addresses: diancrini@uinsby.ac.id (D.C.R. Novitasari), rezzyekocaraka@gmail.com (R.E. Caraka)

Received June 9, 2020

these results, it can be concluded that the feature selection method can improve gained accuracy ±98%.

Keywords: COVID-19; SVM; resnet; GoogleNet; Convolution.

2010 AMS Subject Classification: 62P10, 68T30, 62H25, 62H35.

1. INTRODUCTION

According to the European Center for Disease Prevention and Control website, on April 9, 2020, 1,476,819 people worldwide were tested COVID19 positive and it was stated that 87,816 people died. Coronavirus (Covid-19) originates from the same family as the Middle East Respiratory Syndrome (MERS-CoV) virus first discovered in 2015 and Severe Acute Respiratory Syndrome (SARS-CoV) in 2003 [1]. The initial symptoms often encountered in patients are fever, cough, and myalgia or fatigue [2]. More serious symptoms often experienced by patients are acute respiratory syndromes causing pneumonia and death [3],[4],[5]. Therefore, Coronavirus is also often referred to as acute pneumonia[6]. In line with this, [7] diagnoses made in patients are usually associated with pneumonia and chest X-ray. Along with the rapid development of advanced technology, many studies are conducted on Coronavirus (COVID19) for early detection in patients with this disease using artificial intelligence or also known as machine learning. Artificial intelligence is a modelling system that can study problems like neural network of human brain. Some methods being developed in artificial intelligence are Fuzzy Logic [8], Evolutionary Computing [9], and Machine Learning[10]. Machine Learning is a model approach to a system, so that it can work as closely as possible with neural network of human brain. This method is the most popular one as it is able to study and generalize a problem like human brain. Some applications of this method are prediction and classification. The unique feature of machine learning is training and testing process. One method of machine learning usually used is Neural Network or Artificial Neural Networks. A neural network consists of one or several hidden layers in which there are several neuron units[11]. Neural network method using a hidden multi-layer is called Deep Learning [12]. Deep learning based methodology is usually recommended for the detection of COVID-19 infected patients using X-ray images. This study is conducted to detect whether the patients examined are normal, Coronavirus positive, or just have pneumonia using the convolutional neural network method as feature extraction and support vector machine as classification method or called convolutional support vector machine. This study used different convolutional architectures. There are four stages in this study, that is changes in image size, feature extraction, feature selection, and

classification. Classification is done based on the patient's chest X-ray images. This paper is organized as follows: Section 2 discusses works related to COVID19, Section 2 presents the proposed methodology for COVID19 detection using the Convolutional Support Vector Machine (CSVM) method, Section 3 describes the experiments carried out and presents the results, and finally, Section 4 is about conclusions and future work of this study.

2. PRELIMINARIES

To obtain important information required regarding COVID19 handling , many studies are conducted on COVID19 for early detection in patients with this disease [5],[13],[14] In line with this, [1] is conducted using deep learning, that is inception v3, resnet50, and inception-resnetv2. The best result is obtained on resnet50 with the highest accuracy of 98%, but the study has a limitation of classifying only 2 classes, that is normal and COVID19 based on chest X-ray. Another study conducted by[15] uses deep learning method to classify 3 classes, that is COVID-19, influenza-A viral pneumonia, and normal using CT scan or chest X-ray. The results obtained are quite good, with an accuracy of 89.3%. Because of its improved capabilities, this method is suitable for dealing with complex problems or problems that use large-scale data. As a result, the training process on the deep learning method will require a long time [16]. Another method that can be used aside from deep learning method, which functions to classify image data is the Support Vector Machine (SVM) method [17]. According to previous studies, [18] it yields 100% accuracy for the classification of ADHD (Attention Deficit Hyperactivity Disorder). This shows that SVM is beneficial to be used as a classification of 2 classes and requires a faster time. .In another study conducted by [19] using deep learning method that functions to extract features and support vector machine method that functions as a classification method, obtains the best results also on resnet50 with an accuracy of 95.38% but also can only classify two classes based on chest X-ray. The use of both methods are done to obtain a shorter time and better results.

3. MAIN RESULTS

Main results. The data in this study are chest X-ray images of normal patients, patients infected by Coronavirus, and patients with pneumonia. The data originate from patients who have been tested positive from several countries in the world. Table 1 below shows a breakdown of the data sources used.

Table 1. Data Sources Description

Data	Format	Total	Source
COVID19	.JPEG	102	Github.com
Pneumonia and Normal		204	Kaggle.com

The data were accessed on April 3, 2020, from each website. The data available on the web will be updated regularly by the data provider, so that if the data are accessed on different dates, it will likely to get different amounts of data. Examples of data obtained can be seen in Figure 1.

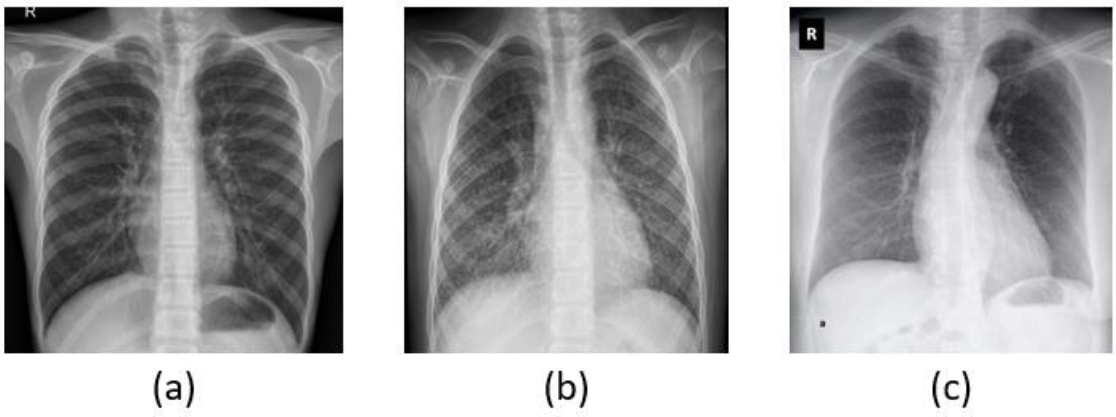

Figure 1. (a) Normal lung; (b) Pneumonia; (c) COVID19

On chest X-ray images of patients with common pneumonia and patients infected by COVID19, they both have abnormalities in the lung parenchyma. However, based on [20], in the chest X-Ray image, the COVID19 patient shows a degree of turbidity in the lungs which is very visible, whereas pneumonia patient has only white patches, but it does not have excess turbidity. There are several types of experiments conducted, that is learning models of GoogleNet, Resnet18, Resnet50, Resnet101 with several feature selections, each using PCA and Relief and class separated. Of the several models that are trained, each of them has a level of accuracy. Figure 2 represents the flow of the experiment.

3.1. Preprocessing

The initial size of the image owned is very diverse, so it is resized to fit the input size of the architecture used, which is 224 × 224. The size selection is based on [21], where the size of the input image in each architecture for the introduced model is 224 × 224. After resizing, the data is divided into training data and testing data. There are 102 data for each class of pneumonia, normal, and COVID19.

3.2.Feature Extraction

Learning in machine learning, after going through the pre-processing stage, is generally continued with feature extraction stage[22],[23]. The method used in this study is convolutional neural network as feature extraction. Convolutional neural network (CNN) is one part of deep learning which is the innovation from the multi-layer perception method and inspired by human artificial neural network[24]. Wiesel and Hubel conducted visual cortex research on the senses of vision of cats [25]. The layer on a convolutional neural network has a 3-dimensional arrangement of length and width, which is the size of the layer and height, which is the depth based on the number of layers. According to the type of layer, [26] represents CNN that can be divided into 2, that is feature extraction layer and fully connected layer. Feature extraction layer, which is located after the input layer in the architectural start, is composed of several layers, each layer is connected to the previous layer. In the feature extraction layer, there are two types of layers, namely convolution layer and pooling layer. Convolutional layers are classified into two (Figure 3), namely convolution layer one dimension used in vector-shaped data such as signal, time series, etc. [27], and convolution layer two dimensions used in two-dimensional data, for example, imagery and others. Pooling layer or subsampling is a reduction in the size of the matrix. There are two types of pooling layers, that is max pooling and average pooling [28].

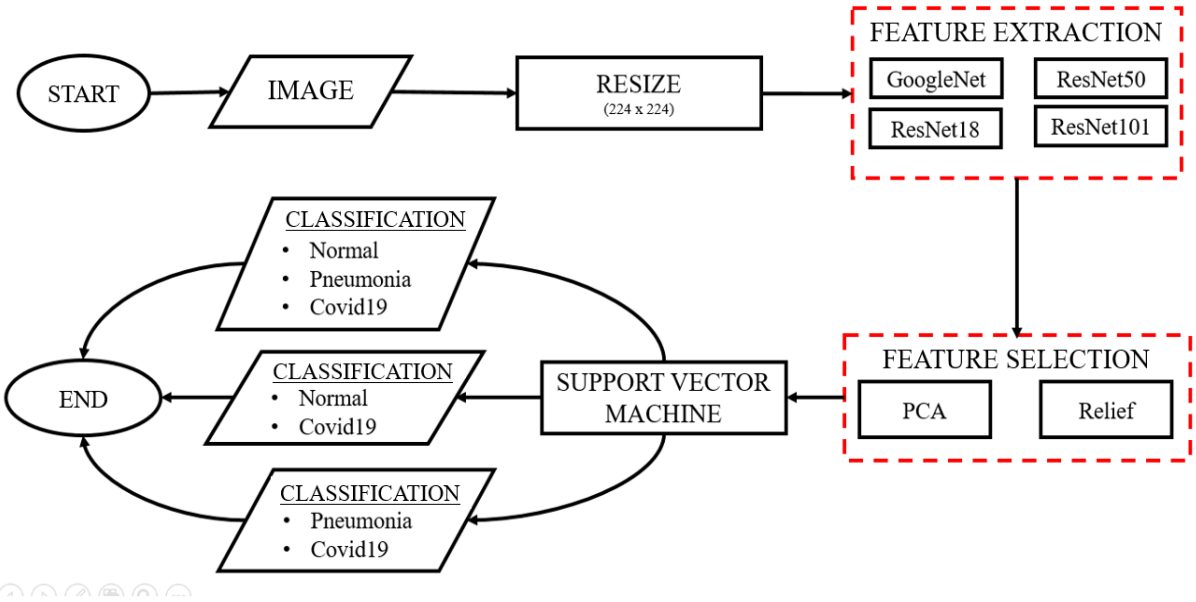

Figure 2. Experiment Flowchart

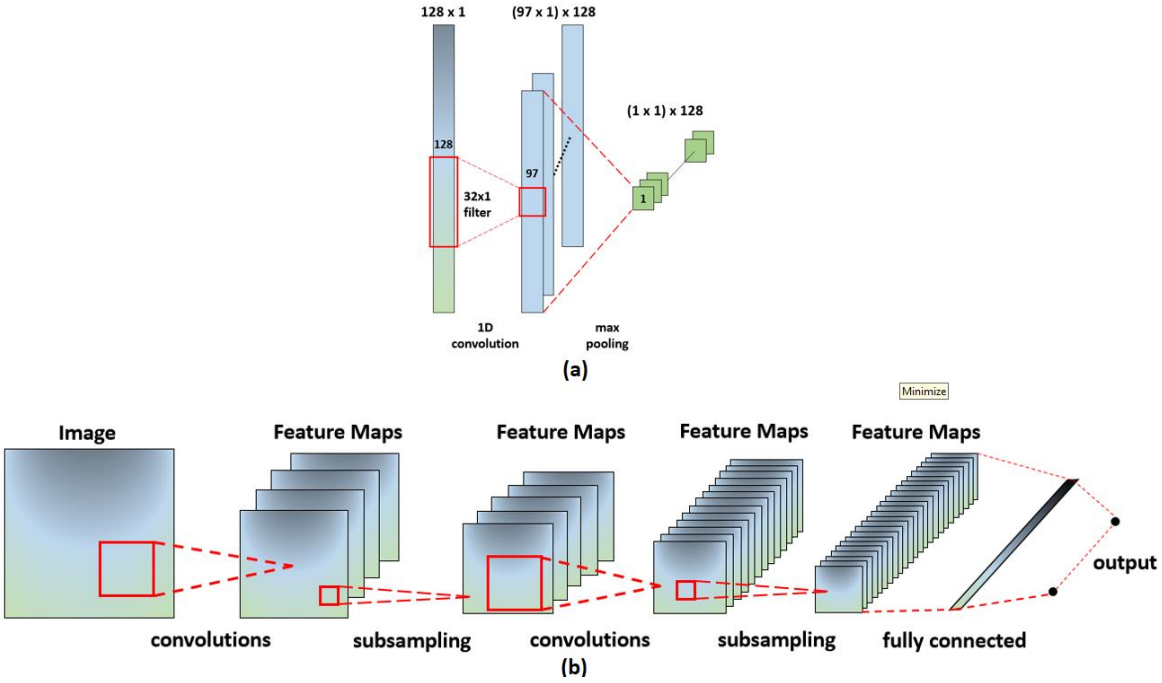

Figure 3. Convolutional 1D Architech [29] (b) *Convolutional 2D* Architech

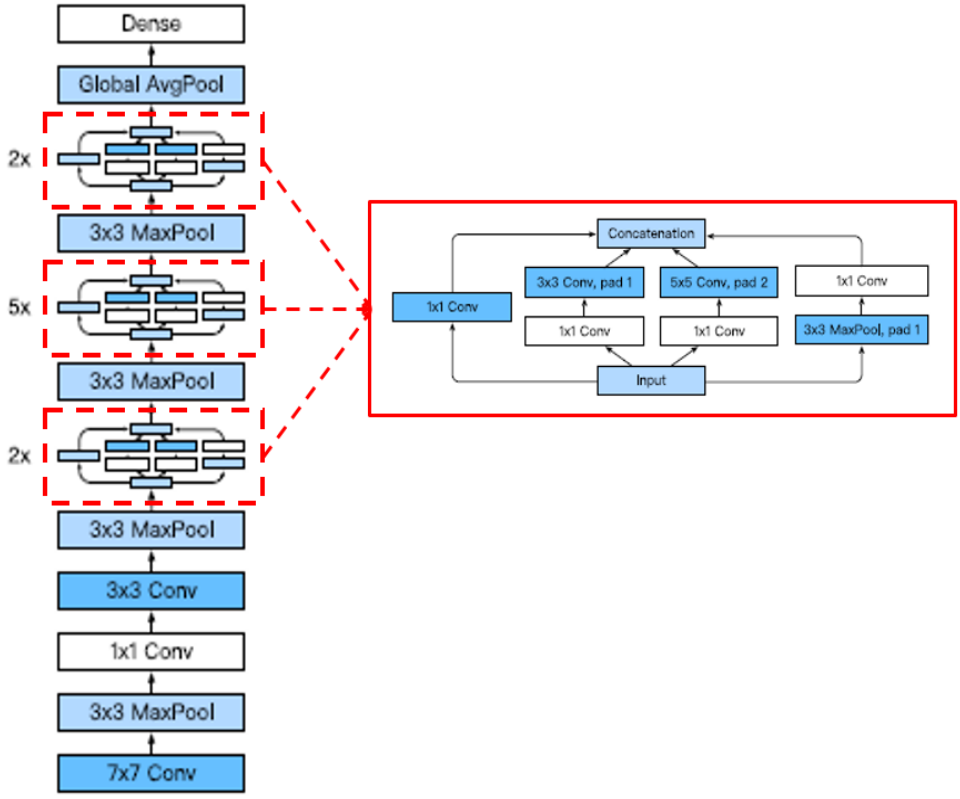

Figure 4. GoogleNet Architech [30]

Fully Connected Layer consists of several layers, and each layer consists of nodes fully connected to the previous layer. The fully connected layer uses a multi-layer perceptron that functions to process data so that the desired results are obtained.

3.2.1 GoogleNet

GoogleNet is one type of architecture of CNN method created by [31]. GoogleNet has inception modules which carry out various convolutions and unify filters for the next layer [32]. The main characteristic of this model architecture is the excellent utilisation of computing resources in the network. GoogleNet architecture can be seen in Figure 4. The core of GoogleNet architecture is that layers in neural networks are extended to the output of various correlation distributions based on the idea that the neural network output of each layer has optimal efficiency if various distributions are done [33].

3.2.1 ResNet

Residual Neural Network (ResNet) is one type of architecture of CNN method created by [21]. ResNet architecture is quite revolutionary as this architecture became state of the art at that time, that is in classification, object detection, and semantic segmentation. The difference between ResNet and other methods is that there are residual blocks as shown in Figure 5. Resnet has various architectures, namely 18, 34, 50, 101 and even up to 152 layers. However, in this study, only ResNet 18, 50, and 101 are used. For the difference in the number of layers it has, it can be seen in Table 2.

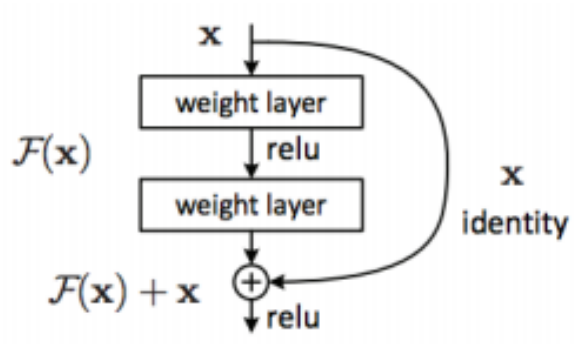

Figure 5. Residual Block of ResNet Architech [21]

The results of each filter of ResNet architecture then go through average pooling before proceeding to the fully connected layer network by using the softmax activation function to

determine the classification results [34].

$$softmax(x)_i = \frac{exp\ (x_i)}{\sum_{j=1}^{n} exp\ (x_j)} \quad (1)$$

Softmax is an activation function that is commonly used to calculate probabilities and carry out multi-class classifications, where softmax values are between 0 to 1 and have a number of 1 if all elements are added using Equation (1) [35]. This function is used at the end of the layer of the fully connected layer used to produce the probability value of an object's function against the existing class. As CNN method only functions as feature extraction, the process carried out only stops at the last layer before entering the fully connected layer as shown in Figure 6. The description of the ResNet's layer arrangement can be seen in Table 2, and that of the google net's can be seen in Figure 4 and each architecture can be seen in Figure 4.

Table 2. Resnet Architech Detail

Layer Name	**Output Size**	**ResNet-18**	**ResNet-50**	**ResNet-101**
conv1	112×112	$7x7, 64, stride\ 2$		
conv2_x	56×56	$3x3\ max\ pool, stride\ 2$		
		$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	
average_pool	1×1	$average\ pool$		
fully_connected	1000	$fully\ connected\ layer$		
softmax		$softmax$		

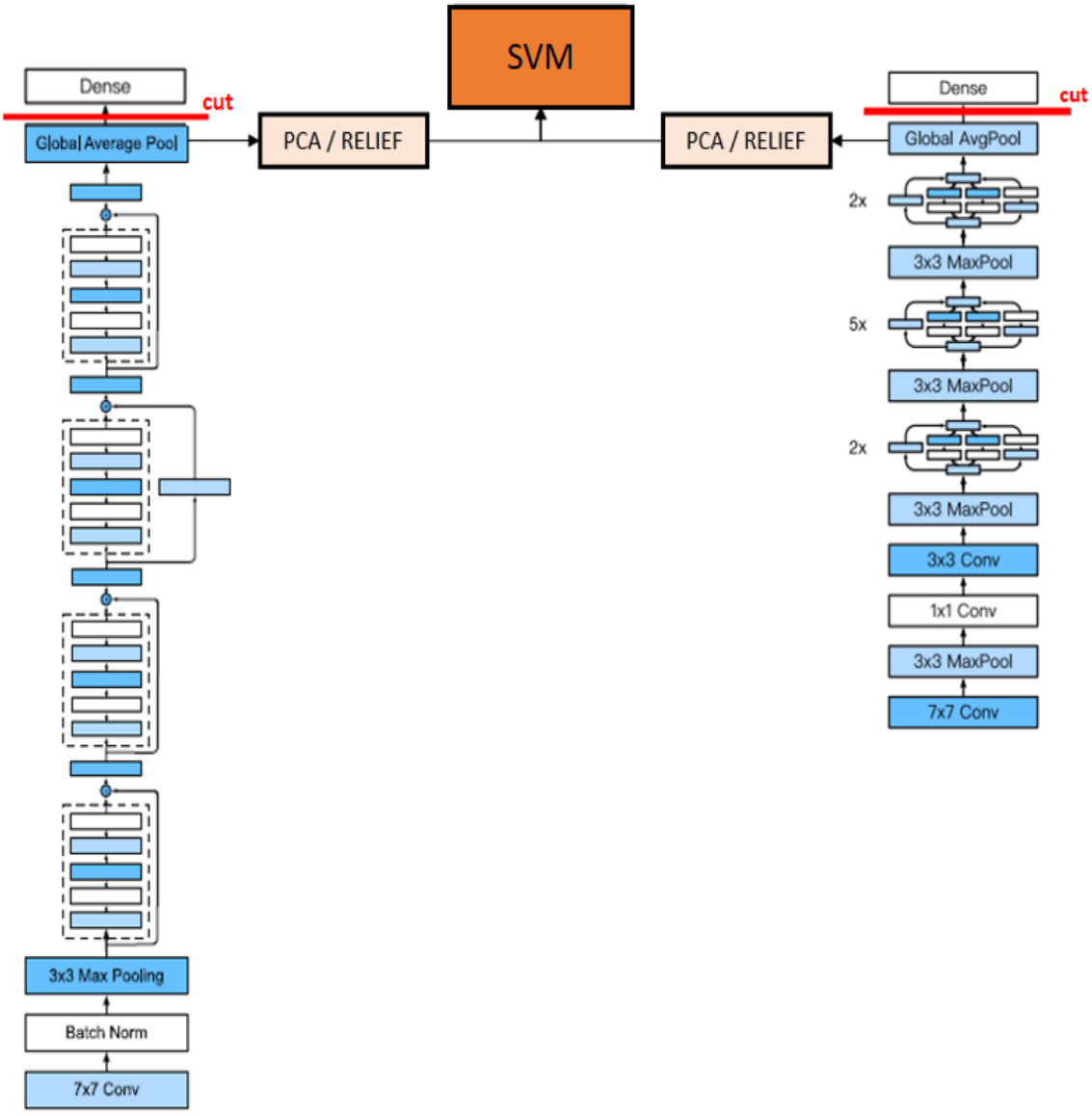

Figure 6. Architectural System Proposed

3.3 Feature Selection

3.3.1 Principal Component Analysis

Principal Component Analysis (PCA) is a method of feature selection that serves to reduce the number of features used without reducing the characteristics of these features [36]. The PCA steps are as follows:

1. Calculate the covariance matrix using Equation (2). However, X and Y are data, and $\bar{X}, \bar{Y}$ represents the average.

$$Cov(X,Y) = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{n-1} \quad (2)$$

2. Calculate the eigenvalue by completing Equation (3). det represents the determinant, *C* is the covariance matrix, *I* is the identity matrix, and λ is the eigenvalue

$$det(C - \lambda I) = 0 \quad (3)$$

3. Calculate the eigenvector by solving Equation (4).

$$(\mathrm{C} - \lambda \mathrm{I})\mathrm{X} = 0 \quad (4)$$

4. Determine the new variable by multiplying the original variable with the eigenvector matrix.

3.3.2 Relief Algorithm

Relief is a feature selection algorithm using weights to measure the effect of these features. The higher the weight, the more important the feature is on the data[37]. This algorithm is inspired by instance-based learning [38]. For example, *S* is training data which has n samples and threshold τ that have values between$0 \leq \tau \leq 1$. This algorithm will detect the effect of the set of features on the target statistically. If it is assumed that each feature is numerical or nominal, then the difference in the value of the feature between the two samples, namely *X* and *Y* is defined by the difference function as follows:

If X_k and Y_k nominal so,

$$diff(X_k, Y_k) = \begin{cases} 0 & , X_k = Y_k \\ 1 & , X_k \neq Y_k \end{cases}$$

If X_k and Y_k numeric so,

$$diff(X_k, Y_k) = \frac{(X_k - Y_k)}{nu_k}$$

nu_k is a normalization unit to make the value of $diff$ be a range of $0 - 1$.

3.4. SVM Classification

Support Vector Machine (SVM) is a method used to classify by finding the best hyperplane value and the results obtained from the optimal classification [17],[39]. Vapnik, Guyon, and Boser first presented this method in 1992 in a workshop called the Annual Workshop on Computational Learning Theory [40]. The primary way of working SVM is linear classification and developments are made to be able to work on non-linear problems[41],[42]. Developments made include adding a kernel trick concept to the method that will be used to find the best hyperplane which can separate the distance (margin) between classes from the data maximally[43]. If the distance between the hyperplane with the closest data from each class is the furthest distance that can be obtained [44][45]. Then the hyperplane can be said to be optimal [46],[47]. Vapnik succeeded in proving that the application of SVM in the real world in the classification problem by separating training data into two classes can work well [48],[49]. The kernels used in SVM are as linear, Radial Basis Function (RBF), sigmoid and polynomial[50].

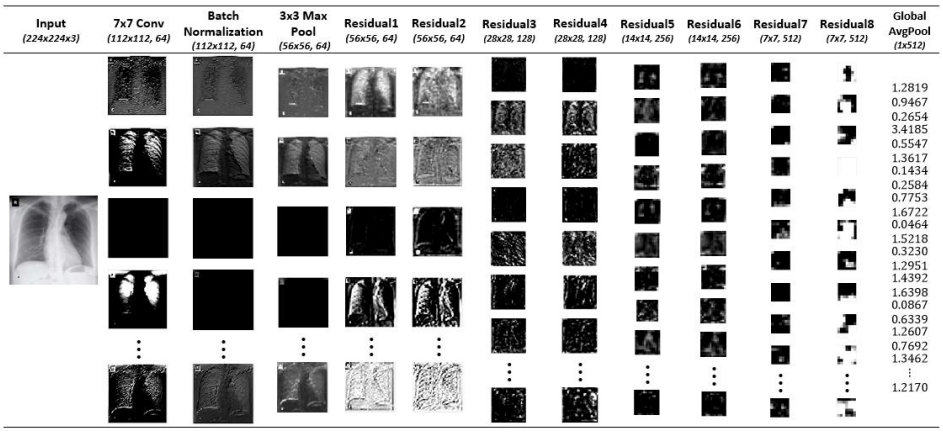

Figure 7. Results of Each Layer of Resnet18

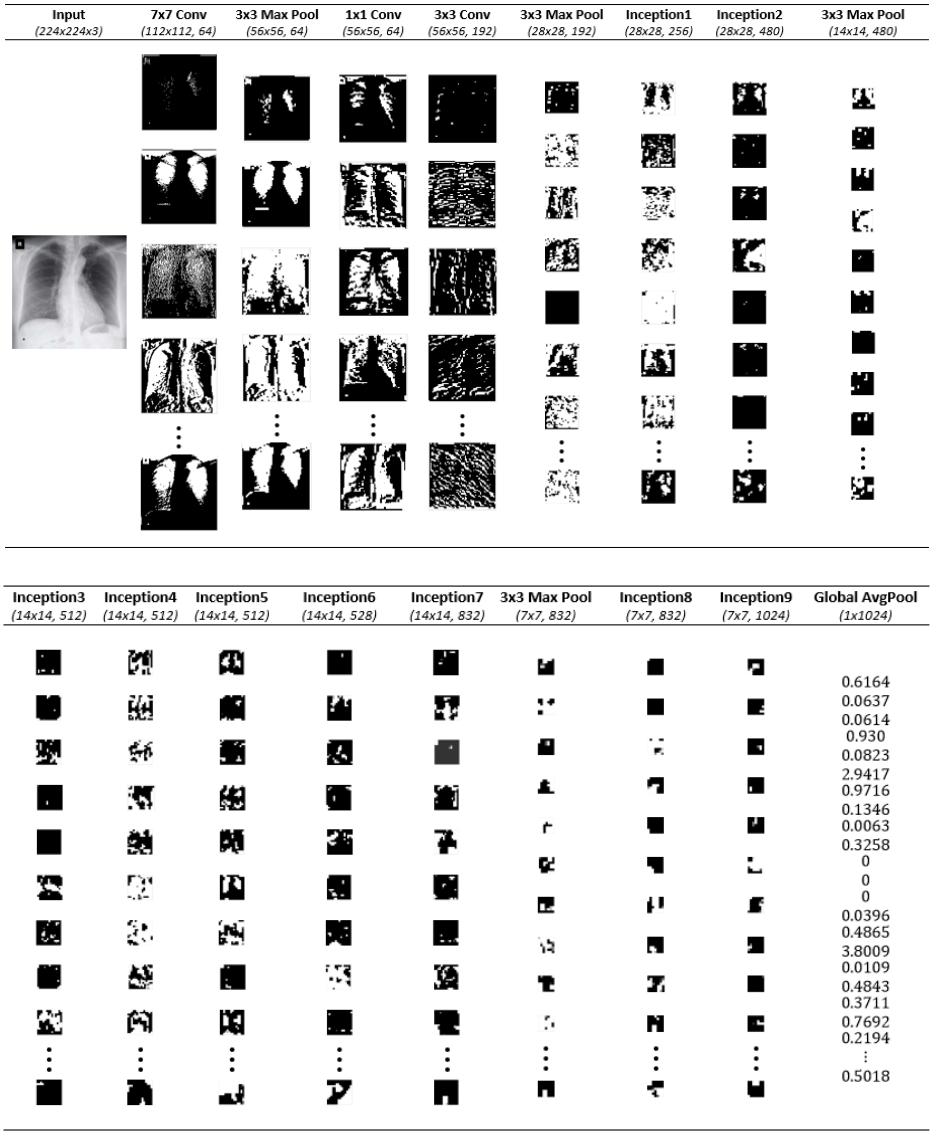

Figure 8. Results of Each Layer of Googlenet

First stage conducted is extracting features of chest X-ray using googlenet, resnet18, resnet50, and resnet101 architectures. Googlnet architecture consists of several convolution layers and 9 pooling layers as well as 9 inception blocks as seen in Figure 4. Results of each layer of googlenet architecture can be seen in Figure 8. Next, resnet18, resnet50, and resnet 101 architectures consist of convolution layer, pooling layer, and several residual layers which have different numbers as shown in Table 2. Resnet18 has 8 residual blocks, resnet50 has16 residual blocks, dan resnet101 has 33 residual blocks. As resnet50 and resnet101 have quite many blocks, only results of resnet18 are shown in Figure 7.The first classification model built can be classified into three classes, that is normal, pneumonia, and Covid19. Based on the experiments carried out, the comparison of results from each model is shown in Table 3.The best accuracy results for the kernel and feature selection methods used are linear kernels and using PCA method. The average accuracy obtained is 91.74%. While the best architecture as feature extraction is resnet50 with an average accuracy of 77%. Nevertheless, overall the best accuracy is 97.33% with the architecture of resnet50 using the polynomial kernel and resnet101 using the sigmoid kernel. Both use the same feature selection method, relief. A confusion matrix of the two can be seen in Figure 9.

Table 3. First Model of Accuration Results

CNN	K-Fold	Linear		Polynomial		Gaussian		Sigmoid		AVERAGE
		PCA	Relief	PCA	Relief	PCA	Relief	PCA	Relief	
Googlenet	Fold 1	88.46	87.18	62.82	85.90	34.62	37.18	82.05	87.18	**70.67**
	Fold 2	85.33	89.33	77.33	88.00	33.33	33.33	86.67	89.33	**72.83**
	Fold 3	92.00	86.67	77.33	85.33	33.33	33.33	85.33	88.00	**72.67**
	Fold 4	93.59	89.74	70.51	88.46	34.62	37.18	92.31	92.31	**74.84**
Resnet18	Fold 1	89.74	87.18	67.95	85.90	34.62	33.33	83.33	91.03	**71.63**
	Fold 2	92.00	88.00	82.67	89.33	33.33	33.33	90.67	89.33	**74.83**
	Fold 3	88.00	96.00	82.67	92.00	33.33	33.33	93.33	92.00	**76.33**
	Fold 4	96.15	93.59	75.64	83.33	34.62	34.62	89.74	89.74	**74.68**
Resnet50	Fold 1	88.46	87.18	76.92	88.46	34.62	33.33	84.62	87.18	**72.60**
	Fold 2	93.33	89.33	85.33	90.67	33.33	33.33	84.00	92.00	**75.17**
	Fold 3	96.00	86.67	82.67	93.33	33.33	33.33	93.33	97.33	**77.00**
	Fold 4	96.15	92.31	71.79	87.18	34.62	37.18	89.74	88.46	**74.68**
Resnet101	Fold 1	93.59	91.03	64.10	89.74	34.62	34.62	84.62	88.46	**72.60**
	Fold 2	89.33	90.67	61.33	97.33	33.33	33.33	89.33	86.67	**72.67**
	Fold 3	94.67	94.67	70.67	94.67	33.33	33.33	90.67	90.67	**75.33**
	Fold 4	91.03	89.74	65.38	83.33	34.62	34.62	88.46	87.18	**71.79**
AVERAGE		**91.74**	**89.96**	**73.45**	**88.94**	**33.97**	**34.29**	**88.01**	**89.80**	

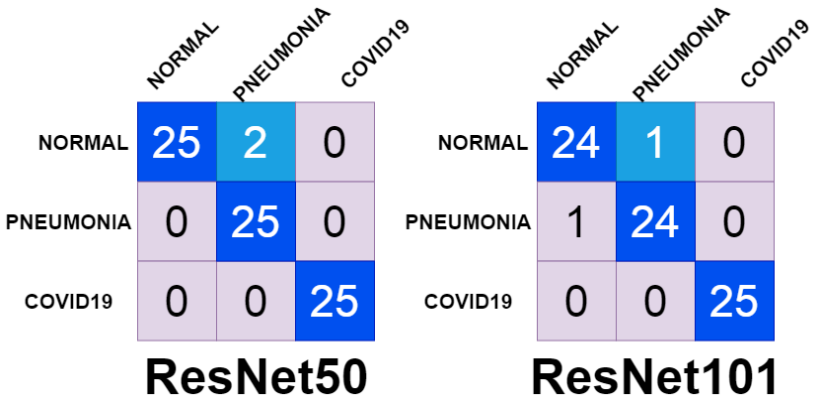

Figure 9. Best Model of Confusion Matrix

The second model is a system that can classify images into two classes, that is normal and COVID19. In line with this, we have compared the results from each of the second models shown in Table 4.The best accuracy results for the kernel and feature selection methods used are linear kernels and using PCA method. The average accuracy obtained is 98.04%. However, our best architecture as feature extraction is resnet50 with an average accuracy of 84.25%. Overall, Table 4 representing our second model is quite good at classifying images into two classes as it manages to obtain quite a lot of accuracy by 100%, particularly using feature selection by PCA method and linear kernel in the SVM process.

Table 4. Second Model of Accuration Results

CNN	K-Fold	Linear		Polynomial		Gaussian		Sigmoid		AVERAGE
		PCA	Relief	PCA	Relief	PCA	Relief	PCA	Relief	
Googlenet	**Fold 1**	98.08	96.15	57.69	94.23	51.92	59.62	94.23	98.08	**81.25**
	Fold 2	96.08	96.08	66.67	98.04	49.02	50.98	98.04	98.04	**81.62**
	Fold 3	92.00	94.00	74.00	96.00	50.00	50.00	94.00	96.00	**80.75**
	Fold 4	98.04	96.08	66.67	96.08	49.02	49.02	94.12	96.08	**80.64**
Resnet18	**Fold 1**	98.08	96.15	69.23	96.15	51.92	51.92	92.31	96.15	**81.49**
	Fold 2	100.00	98.04	60.78	100.00	49.02	49.02	96.08	98.04	**81.37**
	Fold 3	100.00	98.00	86.00	96.00	50.00	50.00	94.00	98.00	**84.00**
	Fold 4	100.00	98.04	64.71	98.04	49.02	49.02	94.12	98.04	**81.37**
Resnet50	**Fold 1**	96.15	96.15	63.46	98.08	51.92	55.77	96.15	94.23	**81.49**
	Fold 2	100.00	96.08	72.55	98.04	50.98	50.98	98.04	92.16	**82.35**
	Fold 3	100.00	100.00	80.00	98.00	50.00	50.00	98.00	98.00	**84.25**
	Fold 4	100.00	100.00	68.63	96.08	49.02	49.02	100.00	100.00	**82.84**
Resnet101	**Fold 1**	98.08	100.00	55.77	100.00	51.92	50.00	96.15	100.00	**81.49**
	Fold 2	98.04	94.12	54.90	96.08	49.02	49.02	98.04	92.16	**78.92**
	Fold 3	100.00	100.00	54.00	98.00	50.00	50.00	98.00	100.00	**81.25**
	Fold 4	94.12	100.00	52.94	96.08	49.02	49.02	92.16	94.12	**78.43**
AVERAGE		**98.04**	**97.43**	**65.50**	**97.18**	**50.11**	**50.84**	**95.84**	**96.82**	

Table 5. Third Model of Accuration Results

CNN	K-Fold	Linear		Polynomial		Gaussian		Sigmoid		AVERAGE
		PCA	Relief	PCA	Relief	PCA	Relief	PCA	Relief	
Googlenet	Fold 1	98.08	96.15	63.46	98.08	50.00	50.00	96.15	96.15	**81.01**
	Fold 2	98.08	98.08	63.46	98.08	50.00	50.00	96.15	98.08	**81.49**
	Fold 3	96.00	96.00	56.00	94.00	50.00	50.00	92.00	96.00	**78.75**
	Fold 4	96.00	96.00	66.00	100.00	50.00	50.00	94.00	90.00	**80.25**
Resnet18	Fold 1	100.00	98.08	90.38	98.08	50.00	50.00	98.08	98.08	**85.34**
	Fold 2	100.00	98.08	71.15	98.08	50.00	50.00	100.00	96.15	**82.93**
	Fold 3	96.00	96.00	80.00	96.00	50.00	50.00	94.00	98.00	**82.50**
	Fold 4	100.00	98.00	88.00	98.00	50.00	50.00	94.00	100.00	**84.75**
Resnet50	Fold 1	98.08	98.08	94.23	96.15	50.00	50.00	96.15	98.08	**85.10**
	Fold 2	100.00	100.00	90.38	100.00	50.00	50.00	96.15	100.00	**85.82**
	Fold 3	98.00	98.00	88.00	98.00	50.00	50.00	96.00	98.00	**84.50**
	Fold 4	100.00	100.00	92.00	100.00	50.00	58.00	96.00	100.00	**87.00**
Resnet101	Fold 1	100.00	100.00	53.85	100.00	50.00	50.00	100.00	100.00	**81.73**
	Fold 2	100.00	100.00	51.92	100.00	50.00	50.00	100.00	100.00	**81.49**
	Fold 3	98.00	98.00	56.00	98.00	50.00	50.00	94.00	98.00	**80.25**
	Fold 4	98.00	100.00	64.00	98.00	50.00	50.00	98.00	100.00	**82.25**
AVERAGE		**98.51**	**98.15**	**73.05**	**98.15**	**50.00**	**50.50**	**96.29**	**97.91**	

In the last experiment, we build a system that can classify images into two classes, that is pneumonia and COVID19. However, the best accuracy results for the kernel and feature selection methods used are linear kernels and using PCA method. Table 5 represents our third model and the average accuracy obtained is 98.51%. The best architecture as feature extraction is resnet50 with an average accuracy of 87%. Overall, the third model is quite good at classifying images into two classes as it manages to obtain quite a lot of accuracy by 100%, particularly using feature selection by PCA method and linear kernel in the SVM process

4. CONCLUSION

Based on several experiments conducted, it can be concluded that the combination of kernel using linear kernel of SVM method and PCA method as a feature selection obtained good results on the three models built. Resnet50 architecture was the best architecture on the three models built. This is in line with the study conducted [19] which also used deep learning methods as feature

extraction and support vector machine methods as a classification method to obtain the best results , also on resnet50 with an accuracy of 95.38% to classify 2 classes based on chest X-ray. In this study, the accuracy was 97.33% for 3 classes and 100% for 2 classes. In a nutshell, we prove that the classification using CNN method is relatively reliable against the parameter changes. By using good and optimal training data, a subset of the training data will also produce a good classification. However, CNN, like other Deep Learning methods, has a weakness, that is time processing. Future studies will compare it with other feature selection techniques such as XGBoost [51], [52], ADABoost or YOLO [53],.

ACKNOWLEDGEMENT

This research was supported by Mathematics Department, Faculty of Science and Technology, Universitas Islam Negeri Sunan Ampel Surabaya, Indonesia and Mathematics Department, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia. This paper in part of MOST-107-2221-E-324 -018-MY2 and MOST-106-2218-E-324 -002 and in part of Chaoyang University of Technology and the Higher Education Sprout Project, Ministry of Education (MOE), Taiwan, under the project name: "The R&D and the cultivation of talent for health-enhancement products".

DATA AVAILABILITY

The analysis datasets used in this paper are available from the corresponding author upon reasonable request.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Dian Candra Rini Novitasari, lead this research and do the Conceptualization, Methodology, Software, Investigation, Data Curation, Formal Analysis, do the writing-original draft preparation, writing-review and editing. Rimuljo Hendradi, supervision, project administration, and funding acquisition. Rezzy Eko Caraka, does the Conceptualization, Methodology, Investigation, writing-original draft preparation, review, editing, and proofing instrument. Yuanita Rachmawati, writing-review and editing. Nurul Zainal Fanani, writing-review and editing. Anang Syarifudin does the

software, writing-review and editing. Toni Toharudin does the supervision, project administration. Rung-Ching Chen does the supervision, project administration, and funding acquisition.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

[1] A. Narin, C. Kaya, Z. Pamuk, Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks. arXiv:2003.10849 [eess.IV]. 2020.

[2] L.L. Ren, Y.M. Wang, Z.Q. Wu et al., Identification of a novel coronavirus causing severe pneumonia in human: a descriptive study, Chin. Med. J. (Engl). 133(9)(2020), 1015-1024.

[3] G. Orive, U. Lertxundi, D. Barcelo, Early SARS-CoV-2 outbreak detection by sewage-based epidemiology, Sci. Total Environ. 732 (2020), 139298.

[4] I. Ghinai et al., First known person-to-person transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in the USA, Lancet, 395 (10230) (2020), 1137–1144.

[5] R. E. Caraka et al., Impact of COVID-19 large scale restriction on environment and economy in Indonesia, Glob. J. Environ. Sci. Manage. 6 (2020), 65–82.

[6] Z.Y. Zu et al., Coronavirus Disease 2019 (COVID-19): A Perspective from China, Radiology. (2020), 200490. https://doi.org/10.1148/radiol.2020200490.

[7] A. Abbas et al., Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network, medRxiv 2020.03.30.20047456; doi: https://doi.org/10.1101/2020.03.30.20047456.

[8] I. Nedeljkovic, Image classification based on fuzzy logic. Remote Sens. Spat. Inf. Sci. 34 (2006), 1–6.

[9] S. Cagnoni, ed., Genetic and evolutionary computation for image processing and analysis, Hindawi Publ, New York, NY, 2007.

[10] L. Palagi, A. Pesyridis, E. Sciubba, L. Tocci, Machine Learning for the prediction of the dynamic behavior of a small scale ORC system, Energy. 166 (2019), 72–82.

[11] R. E. Caraka, S. A. Bakar, M. Tahmid, H. Yasin, and I. D. Kurniawan, Neurocomputing Fundamental Climate Analysis, Telkomnika, 17 (4) (2019), 1818–1827.

[12] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, Deep learning for visual understanding: A review, Neurocomputing, 187 (2016), 27–48.

[13] T. Toharudin et al., Bayesian Poisson Model for COVID-19 in West Java Indonesia, Sylwan, 164 (2020), 03c2j.

[14] R. S. Pontoh, Z. Solichatus, Y. Hidayat, R. Aldella, N. M. Jiwani, and Sukono, Covid-19 modelling in south korea using a time series approach, Int. J. Adv. Sci. Technol. 29 (7) (2020), 1620–1632.

[15] C. Butt, J. Gill, D. Chun, B.A. Babu, Deep learning system to screen coronavirus disease 2019 pneumonia, Appl Intell. (2020). https://doi.org/10.1007/s10489-020-01714-3.

[16] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, Commun. ACM. 60 (2017), 84–90.

[17] A. Z. Foeady, D. C. R. Novitasari, A. H. Asyhar, Diabetic Retinopathy : Identification and Classification using Different Kernel on Support Vector Machine, Proc. Int. Conf. Math. Islam, 2018, vol. 2019, 129–134.

[18] J.C. Bledsoe, D. Xiao, A. Chaovalitwongse, S. Mehta, T.J. Grabowski, M. Semrud-Clikeman, S. Pliszka, D. Breiger, Diagnostic Classification of ADHD Versus Control: Support Vector Machine Classification Using Brief Neuropsychological Assessment, J. Atten. Disord. (2016), doi: 10.1177/1087054716649666.

[19] P. Kumar and S. Kumari, Detection of coronavirus Disease (COVID-19) based on Deep Features, Preprints 2020, 2020030300, doi: 10.20944/preprints202003.0300.v1.

[20] S.H. Yoon, K.H. Lee, J.Y. Kim, Y.K. Lee, H. Ko, K.H. Kim, C.M. Park, Y.-H. Kim, Chest Radiographic and CT Findings of the 2019 Novel Coronavirus Disease (COVID-19): Analysis of Nine Patients Treated in Korea, Korean J Radiol. 21 (2020) 494.

[21] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 2016 (2016), 770–778.

[22] R. E. Caraka, R. C. Chen, H. Yasin, B. Pardamean, T. Toharudin, S. H. Wu, Prediction of Status Particulate Matter 2.5 using State Markov Chain Stochastic Process and HYBRID VAR-NN-PSO, IEEE Access, 7 (2019), 161654–161665.

[23] A.H. Asyhar, A.Z. Foeady, M. Thohir, A.Z. Arifin, D.Z. Haq, D.C.R. Novitasari, Implementation LSTM Algorithm for Cervical Cancer using Colposcopy Data, in: 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), IEEE, Fukuoka, Japan, 2020: pp. 485–489.

[24] C. Dewi, R.-C. Chen, T. Shao-Kuo, Evaluation of Robust Spatial Pyramid Pooling Based on Convolutional Neural Network for Traffic Sign Recognition System, Electronics, 9 (6) (2020), 889.

[25] D.H. Hubel, T.N. Wiesel, Receptive fields of single neurones in the cat's striate cortex, J. Physiol. 148 (1959), 574–591.

[26] S. Dutta, B. C. S. Manideep, S. M. Basha, R. D. Caytiles, and N. C. S. N. Iyengar, Classification of diabetic retinopathy images by using deep learning models, Int. J. Grid. Distrib. Comput. 11 (2018), 89–106.

[27] Y. Kim, Convolutional Neural Networks for Sentence Classification, in: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Association for Computational Linguistics, Doha,

Qatar, 2014: pp. 1746–1751.

[28] F. Moutarde, “Deep-Learning: general principles + Convolutional Neural Networks,” no. March, 2018. http://perso.mines-paristech.fr/fabien.moutarde/ES_LSML/Slides/deepLearning-convNets_course-FabienMOUTARDE_2pp.pdf

[29] H. Lim, J. Park, K. Lee, and Y. Han, Rare Sound Event Detection Using 1d Convolutional Recurrent Neural Networks, Detection and Classification of Acoustic Scenes and Events 2017, https://www.cs.tut.fi/sgn/arg/dcase2017/documents/workshop_presentations/DCASE2017Workshop_Lim_205_poster.pdf.

[30] A. Goleman, Daniel; Boyatzis, Richard; Mckee, Dive into Deep Learning, J. Chem. Inf. Model. 53 (9) (2019), 1689–1699.

[31] C. Szegedy, et al., Going deeper with convolutions, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Boston, MA, USA, 2015: pp. 1–9.

[32] A.Z. Foeady, D.C.R. Novitasari, A.H. Asyhar, Diabetic Retinopathy: Identification and Classification using Different Kernel on Support Vector Machine:, in: Proceedings of the International Conference on Mathematics and Islam, SCITEPRESS - Science and Technology Publications, Mataram, Indonesia, 2018: pp. 72–79.

[33] J.-H. Kim, S.-Y. Seo, C.-G. Song, K.-S. Kim, Assessment of Electrocardiogram Rhythms by GoogLeNet Deep Neural Network Architecture, J. Healthcare Eng. 2019 (2019), 2826901.

[34] P. Napoletano, F. Piccoli, R. Schettini, Anomaly Detection in Nanofibrous Materials by CNN-Based Self-Similarity, Sensors. 18 (2018), 209.

[35] J.M. Czum, Dive Into Deep Learning, J. Amer. Coll. Radiol. 17 (2020), 637–638.

[36] J. M. Kinser, J. M. Kinser, Principle Component Analysis, in Image Operators, 2018.

[37] D.C. Rini Novitasari, A.Z. Foeady, M. Thohir, A.Z. Arifin, K. Niam, A.H. Asyhar, Automatic Approach for Cervical Cancer Detection Based on Deep Belief Network (DBN) Using Colposcopy Data, in: 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), IEEE, Fukuoka, Japan, 2020: pp. 415–420.

[38] C. Cortes, V. Vapnik, Support-Vector Networks, Mach. Learn. 20 (3) (1995), 273–297.

[39] D.D. Prastyo, H.A. Khoiri, S.W. Purnami, Suhartono, S.-F. Fam, N. Suhermi, Survival Support Vector Machines: A Simulation Study and Its Health-Related Application, in: M.W. Berry, A. Mohamed, B.W. Yap (Eds.), Supervised and Unsupervised Learning for Data Science, Springer International Publishing, Cham, 2020: pp. 85–100.

[40] W. Haerdle, D. D. Prastyo, C. M. Hafner, Support Vector Machines with Evolutionary Feature Selection for Default Prediction, in Oxford Handbook of Applied Nonparametric and Semiparametric Econometrics and

Statistics, Oxford University Press, 2014.

[41] R. C. Chen, C. H. Hsieh, Web page classification based on a support vector machine using a weighted vote schema, Expert Syst. Appl. 31 (2) (2006), 427–435.

[42] D.C.R. Novitasari, et al., Whirlwind Classification with Imbalanced Upper Air Data Handling using SMOTE Algorithm and SVM Classifier, J. Phys.: Conf. Ser. 1501 (2020), 012010.

[43] R. E. Caraka, S. A. Bakar, Evaluation Performance of Hybrid Localized Multi Kernel SVR (LMKSVR) In Electrical Load Data Using 4 Different Optimizations, J. Eng. Appl. Sci. 13 (17) (2018), 7440-7449.

[44] C. Dewi, R. C. Chen, Random Forest and Support Vector Machine on Features Selection for Regression Analysis, Int. J. Innov. Comput. Inf. Control, 15 (6) (2019), 2027–2037.

[45] M. Thohir, A.Z. Foeady, D.C.R. Novitasari, A.Z. Arifin, B.Y. Phiadelvira, A.H. Asyhar, Classification of Colposcopy Data Using GLCM-SVM on Cervical Cancer, in: 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), IEEE, Fukuoka, Japan, 2020: pp. 373–378.

[46] H. Yasin, R. E. Caraka, A. Hoyyi, and Sugito, Stock price modeling using localized multiple kernel learning support vector machine, ICIC Express Lett. Part B Appl. 11 (4) (2020), 333–339.

[47] D.C.R. Novitasari, et al. Weather Parameters Forecasting as Variables for Rainfall Prediction using Adaptive Neuro Fuzzy Inference System (ANFIS) and Support Vector Regression (SVR), J. Phys.: Conf. Ser. 1501 (2020), 012012.

[48] R.E. Caraka, S.A. Bakar, M. Tahmid, Rainfall forecasting multi kernel support vector regression seasonal autoregressive integrated moving average (MKSVR-SARIMA), in: Selangor, Malaysia, 2019: p. 020014.

[49] D.D. Prastyo, H.A. Khoiri, S.W. Purnami, Suhartono, S.-F. Fam, Simulation Study of Feature Selection on Survival Least Square Support Vector Machines with Application to Health Data, in: B.W. Yap, A.H. Mohamed, M.W. Berry (Eds.), Soft Computing in Data Science, Springer Singapore, Singapore, 2019: pp. 34–45.

[50] R. E. Caraka, R. C. Chen, T. Toharudin, M. Tahmid, B. Pardamean, and R. M. Putra, Evaluation Performance of SVR Genetic Algorithm and Hybrid PSO in Rainfall Forecasting, ICIC Express Lett. Part B Appl. 11 (7) (2020), 631–639.

[51] D. Nielsen, Tree Boosting With XGBoost-Why Does XGBoost Win “Every” Machine Learning Competition? Master’s Thesis, Norwegian University of Science and Technology, Trondheim, Norway, 2016.

[52] R. C. Chen et al., An End to End of Scalable Tree Boosting System, Sylwan, 165 (1) (2020), 1–11.

[53] J. Redmon, A. Farhadi, YOLOv3: An Incremental Improvement, ArXiv:1804.02767 [Cs]. (2018).