Available online at http://scik.org

Commun. Math. Biol. Neurosci. 2020, 2020:85

https://doi.org/10.28919/cmbn/5035

ISSN: 2052-2541

MYCOBACTERIUM TUBERCULOSIS IMAGES CLASSIFICATION BASED ON COMBINING OF CONVOLUTIONAL NEURAL NETWORK AND SUPPORT VECTOR MACHINE

AERI RACHMAD[1], NUR CHAMIDAH[2,*], RIRIES RULANINGTYAS[3]

[1]Department of Informatics, Faculty of Engineering, University of Trunojoyo Madura, Bangkalan 69162, Indonesia

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

[3]Department of Physics, Faculty of Science and Technology, Airlangga University, Surabaya 60115, Indonesia

Abstract: Mycobacterium Tuberculosis (TB bacteria) is a rod-shaped bacterium with a very small size. This bacterium can cause lung disease known as Tuberculosis. These TB bacteria can be seen at least by using a conventional microscope with magnification 1000 times. Images that have been seen in a microscope will be further processed by digital image processing. The data used in this study were 100 captions. Based on the color of the TB bacteria, a sputum image is detected and then cropping is done. Total data on TB bacteria and non-bacterial crops in automatic cropping were 1266 crops consisting of 633 TB bacteria and 633 non-TB bacteria. The size of the TB bacteria and open TB bacteria have different pixel sizes, so it needs to resize the image with a size of 50 x 50 pixels. There are several Convolutional Neural Networks (CNN) architectures that have been tried in solving classification problems among them LeNet, AlexNet, ZFNet, GoogleNet, VGGNet and ResNet. In other studies, the accuracy was 95.05% using the Inception V3 method. In the case of this classification of TB bacteria, researchers proposed the

---

[*]Corresponding author

E-mail address: nur-c@fst.unair.ac.id

Received September 19, 2020

ResNet-101 architecture with 224x224x3 pixel input data specifications, 347layer and 1000 full connected layer (fc1000). As for the classification, researchers used the Support Vector Machine (SVM) to determine TB bacteria or not TB bacteria. The results of this study resulted in an accuracy of 97.6%, 97.9% precision, 97.4% recall and F1 score 97.6%.

Keywords: mycobacterium tuberculosis; conventional microscope; CNN; support vector machine.

2010 AMS Subject Classification: 68U10, 68T10, 62H35.

1. INTRODUCTION

Tuberculosis is one of the biggest causes of death in the world. Every year, there are around 10 million people who are infected with tuberculosis [1]. There are three biggest countries that have tuberculosis sufferers namely India (23%), Indonesia (10%) and China (10%) of the total number of tuberculosis sufferers in the world [2]. This disease attacks humans regardless of age, both children and adults. In children - TB disease can attack children aged 0 to 14 years. From the WHO data in 2016 the death of a child died [3]. But usually 90% of people with tuberculosis attack adults. While the ratio of adults is between men and women, men are two times the number of women [2]. This tuberculosis disease not only attacks the lung organs, but this disease can attack other organs such as brain, kidneys, skin and bones [4]. TB disease can be transmitted through the air media when TB sufferers are coughing, spitting, and sneezing [3].

The tuberculosis bacteria has a small calf. These bacteria have varying sizes. i.e. lengths range from 1 to 4 µm with thicknesses ranging from 0.3 to 0.6 µm [5]. Tuberculosis bacteria can be seen using a conventional microscope with 1000x magnification. TB bacteria found in sputum have a unique characteristic that is reddish in color which is called acid-resistant bacteria (BTA). This reddish color appears due to staining fluid using Ziehl-Neelsen (ZN) staining methods [6].

There are several studies relating to the identification and classification of sputum image which contain TB bacteria. In 2012, Purwanti and Widiyanti used the learning vector quantization (LVQ) method with an accuracy of 91.33% [7], while other researchers produced an accuracy of 77.5%, using artificial neural networks [8]. In 2018 Mithra and Emmanuel obtained an accuracy of 91.38% using the Gaussian Fuzzy Neural Network (GFNN) method [6]. While Dinesh obtained an accuracy of 95.05% using the deep neural networks method [9].

CNN method has been widely used by researchers before. This is because this algorithm has been claimed as the best method in solving the problem of acquiring an object [10]. several CNN architectural models proposed in previous studies such as LeNet, AlexNet, ZFNet, GoogleNet, VGGNet and ResNet [11]. CNN architecture using ResNet can classify images on ImageNet dataset images with an accuracy of 80.62% for Top-1 and has won several competitions in 2015 by beating the architecture of GoogleNet, AlexNet and VGGNet [12].

Based on the literature study above, the researchers developed a tuberculosis bacterial classification system using the CNN Architecture which was used using ResNet-101 in combination with SVM as its classifier. In this study the targets achieved were to assist clinical pathologists in completing their task of identifying tuberculosis bacteria and producing the best results of classification.

2. PRELIMINARIES

In Fig. 1, it shows that there are several steps involved in this research on processing TB images. The first stage, the sputum image data is detected based on the color characteristics of the TB bacteria then cropping it according to the size of the TB bacteria. The second stage, all TB or non-TB bacteria that have been detected will be divided into two classes, then resize the image to a size of 50 x 50 pixels. The third stage, all bacteria will be classified using CNN. The CNN architecture used in this study is ResNet-101. After that it is classified using SVM to recognize TB or non-TB bacteria.

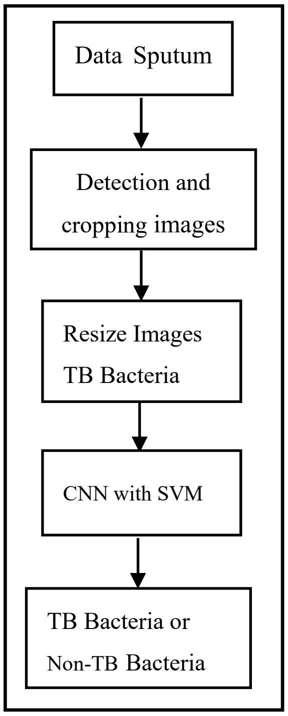


Fig.1 Design System

2.1. Data Input

The image data used in this study is a 800 x 600 pixel sputum image which has been given Ziehl- Neelsen (ZN) staining methods. The sputum image was taken using a Labomed Digi 3 digital microscope equipped with Lx400 and an iVu 5100 digital camera module with a 5.0 megapixel capacity. This sputum image has a resolution of 120 with a depth of 24 bits and an enlargement of 100x [13]. The following is the sputum image shown in Fig. 2.

Sputum image data were taken as many as 50 images. Out of those 50 images, then the images of individual TB bacteria were taken. The segmentation technique uses tolls software to separate bacteria in sputum image images. The dimensions of TB bacterial images vary depending on the size of the bacteria and the position of the bacteria as shown in Fig. 3.

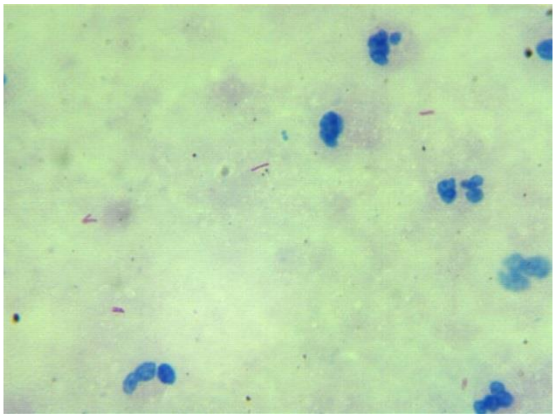


Fig. 2 Sputum image with 100x magnification

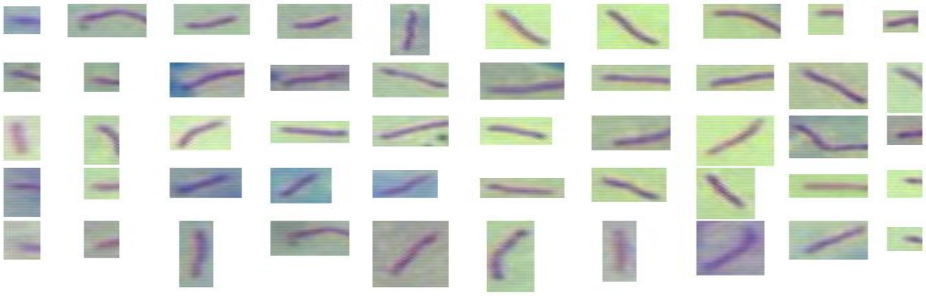


Fig. 3 Single TB bacteria

2.2. Data Augmentation

Data augmentation is one way to increase the amount of data by adding variants to the initial data. There are several ways of data augmentation such as geometric transformation, kernel filter and color modification. Geometric transformation is an augmentation method that is done by changing the position of the image pixels such as translation, rotation, dilation, and affine transformation. But in general, the basic geometric transformations that are used are reflection, rotation and dilation [14].

2.3. Convolusional Neural Networks (CNN)

CNN is a feed-forward neural network developed from the Perseptron Multilayer algorithm. Each set of parameters to be trained in the convolutional layer (referred to as a convolutional filter) extracts meaningful visual concepts from the original input image; while the set of parameters to be trained in a connected layer fully classifies the visual features extracted into the target class such as TB bacterial forms or Not TB bacteria. Convolutional layers are abstract hierarchies of visual concepts from raw input images so that the lower convolutional layers extract low-level features such as colors or shapes, while the higher ones extract high-level visual concepts such as sub-parts of an object [15].

CNN algorithm is an algorithm that is widely used by researchers to observe an object and make an introduction to an object [15]. In Fig. 4 shows the input data that enters the CNN through the Convolutional Layer then is processed to the Rectified Linear Unit (ReLU) layer followed by the Pooling Layer process. The end of this process is the fully Connected layer and Softmax to obtain the probability value of data for each class, where input data are classified into the class with the highest probability [10].

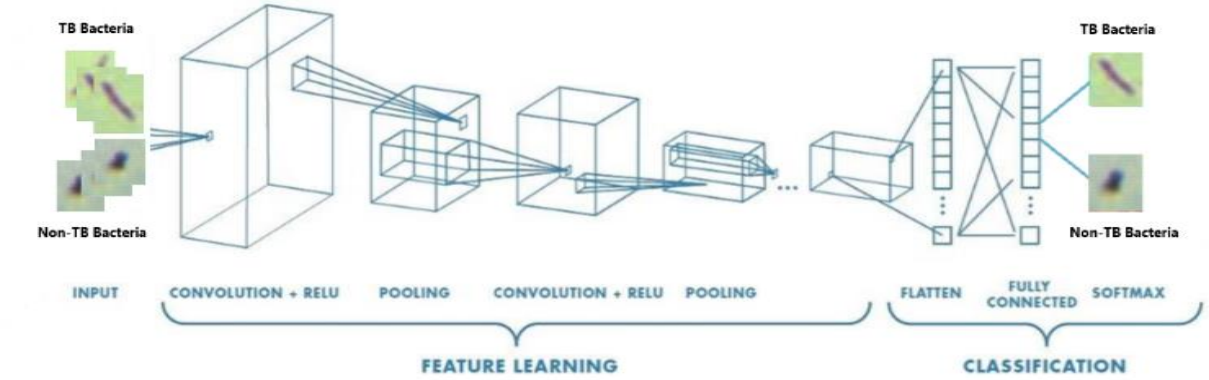


Fig. 4 CNN architecture in general

2.4. Deep Residual Network (ResNet)

Deep Residual Network is one of the CNN network architectures that does residual learning, where residual learning skips the connection from a layer to the input of the deep layer, as shown in Fig. 5. This method was summarized by He et al [12].

This ResNet architecture was built to fix problems on CNN, this is because CNN takes a long time and is limited to a certain number of layers. This ResNet can be used to overcome this problem by applying skip connections or shortcuts. Fig. 6 shows the skip connection on ResNet located on 2 to 3 layers containing ReLU and batch normalization between architectures. Besides, the advantages of ResNet compared to other CNN methods is that the performance of ResNet does not decrease even though the architecture is getting deeper [12].

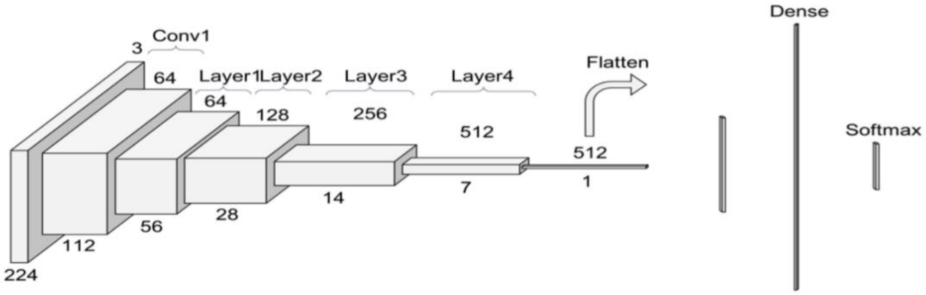


Fig. 5 Architecture Residual Network

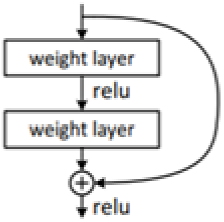


Fig. 6 Deep Residual Network

The architecture of ResNet has several kinds of architecture based on the number of layers, namely ResNet-18 with 18 layers, ResNet -34 with 34 layers, ResNet-50 with layer 50, ResNet-101 with 101 layers and ResNet-152 with 152 layers [12].

2.5. Support Vector Machine

The Support Vector Machine (SVM) method is a classification method that can solve linear and non-linear data problems. This method was first introduced by Vapnik in 1995 which was used to make predictions in both regression and classification cases. The classification using the SVM method is divided into two processes, namely the training process and the testing process. SVM training process uses an input which is resulted from a feature matrix extraction process. Whereas the SVM testing utilizes the projection matrix produced in the feature extraction process which is then multiplied by the test data (test sample) as input. SVM has work principles namely structural risk minimization (SRM) [16].

The workings of the SVM method are to train a set of classification data with an algorithm to produce a classification model that can help in predicting the categories of new data. SVM has many advantages in terms of classification. SVM is one of the leading methods of machine learning that has good results in terms of classification and prediction [17].

2.6. Confusion Matrix

From testing the training data and testing data it produces a confusion matrix as shown in Table 1.

Table. 1 Confusion matrix

		Actual Value	
		True	False
Predicted Value	True	TP (True Positive)	FP (False Positive)
	False	FN (False Negative)	TN (True Negative)

There are several methods to measure the performance of a classification model, namely by finding the accuracy value, Precision, Recall and F1 score of a model [18].

a. Accuracy is a measurement method to calculate the correct prediction value that has been done by the classifier. This accuracy value is obtained by adding up the correct classification results i.e. TP and TN divided by the total amount of test data

$$\mathrm{acc} = \frac{\mathrm{TP+TN}}{\mathrm{TP+FP+TN+FN}} \tag{1}$$

b. Precision is calculated based on the results of the correct predictive value (TP) compared to the actual TP and FP values. The function of Precision is to measure the correct prediction results of all predictions in a positive class.

$$\mathrm{p} = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \tag{2}$$

c. Recall serves to measure positive predictions that are true (TP) on all true predictions.

$$R = \frac{TP}{TP + FN} \tag{3}$$

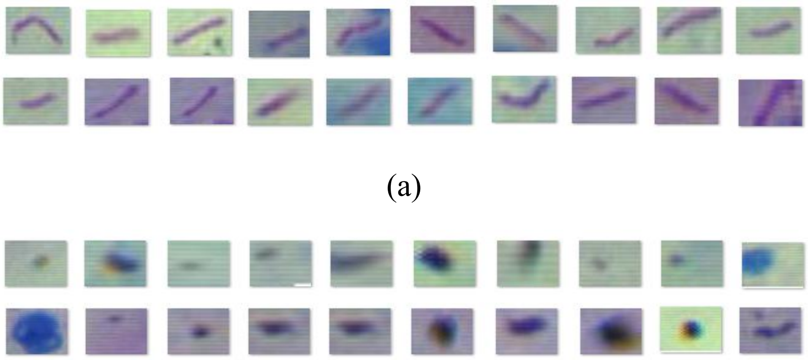
d. F1-Score is 2 times the value of precision and recall value divided by the sum.

$$F1Score = \frac{2 * p * r}{p + r} \tag{4}$$

3. MAIN RESULTS

In general, the implementation of this research uses hardware in the form of a laptop with Windows 10 with 64 bits as the operating system. The detailed specifications of the laptop include: Intel Core i5 Processor 1.6 - 1.8 GHz, 8 GB RAM. As for the software used in this study is MATLAB R2019a.

The trial in this study used 50 sputum image samples with a size of 800 x 600 pixels that have been given Ziehl-Neelsen staining. The number of bacteria found in 50 sputum images was 1266 images consisting of 633 images of TB bacteria and 633 images of non-TB bacteria. The size of the bacteria varies and the number of bacteria in each sputum image is different. Therefore, it is necessary to do a 50 x 50 pixel resize process so that the image has the same size, this is as shown in Fig. 7.



(a)

(b)

Fig. 7 (a) Image of 50 x 50 pixel TB bacteria and (b) Image of 50 x 50 pixel Non-TB bacteria

In the trial process, the data was divided into 2 classes, namely TB Bacteria and Non-TB Bacteria classes. The distribution of the trial process is 70% of the total data of 866 data for training data as much as 30% of the total data of 380 for testing data. They were tested 5 times using the CNN ResNet-101 method using SVM as the classifier.

The results obtained from these trials can be seen in Fig. 8. The highest value in precision is in the 3rd trial at 97.9%. While the highest value in recall is in the 3rd trial and the 5th trial with a value of 97.4%, for the highest accuracy and F1 Score there are in the 3rd trial with values, 97.6% and 97.6%.

In addition to testing 5 times, they were also tested with different classifiers namely KNN (K-Nearest Neighbor), Naïve Bayes and Decision Tree. As shown in Table 2, the SVM classifier test results have the highest accuracy value with a value of 96.56% and the lowest accuracy value in the KNN and decision tree classifier with a value of 85.32%. For Naïve Bayes precision values have the lowest value of 78.76% and the highest use SVM with a value of 96.96%. The highest recall value remains in the SVM method with a value of 96.28% and decision tree is the lowest classifier with a value of 86.58%. While the highest F1 score uses SVM with a value of 96.6% and the lowest is on Naïve Bayes with a value of 84.28%.

Table. 2 The average value using Resnet-101 with four classifiers

Measurement Average	**SVM**	**KNN**	**Naïve Bayes**	**Decision Tree**
Accuracy	96.56	91.44	85.32	85.32
Precision	96.96	92.86	78.76	83.58
Recall	96.28	90.34	90.66	86.58
F1 Score	96.6	91.56	84.28	85.02

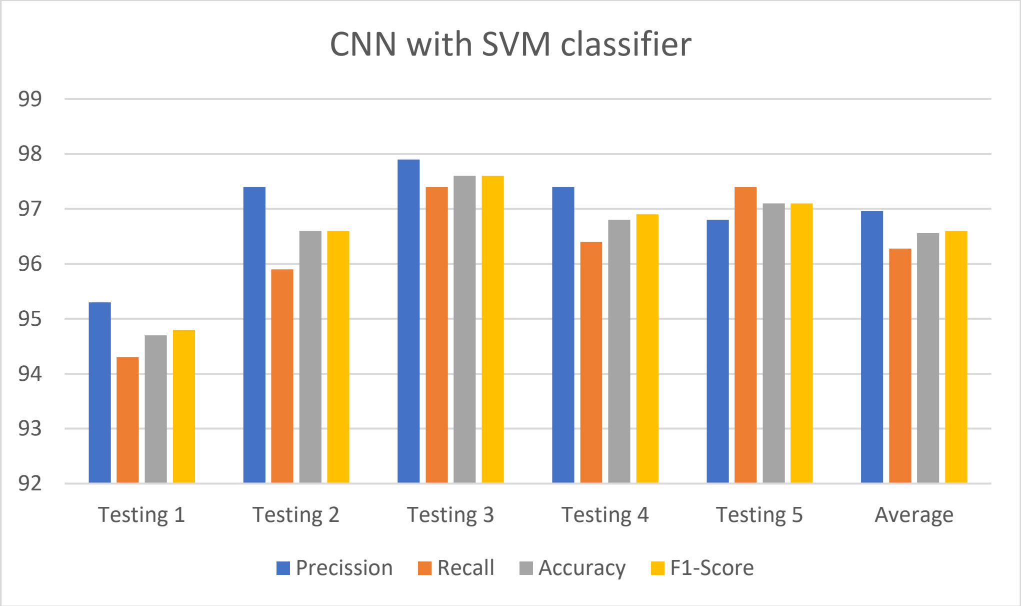


Fig. 8 Testing 5 times propose method

4. CONCLUSIONS

The results of the 5 experiments above show that the average value of accuracy using CNN with ResNet-101 architecture with SVM classifier is 96.56%. With the highest accuracy value of 97.6%. This shows that the SVM classifier is suitable to be used on CNN with ResNet architecture.

The use of classifiers in the classification of TB bacteria depends on the classifier method that is compatible with the CNN method with the SVM classifier. Thus, the KNN, Naïve Bayes and Decision Tree classifier are not appropriate to be used as a classifier to be combined with ResNet-101.

ACKNOWLEDGEMENTS

Authors thank to Director of the Directorate of Research and Public Service, the Directorate General Reinforcing of Research and Development, the Ministry of Research, Technology, and Higher Education of the Republic of Indonesia for financial support of this research through the doctoral dissertation research grant in the fiscal year 2020 with Contract Number: 821/UN3.14/PT/2020.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

[1] World Health Organization, Global tuberculosis report 2019. Geneva (Switzerland): World Health Organization, (2019), https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-report-2019

[2] Ministry of Health of Republic of Indonesia. Profil Kesehatan Republik Indonesia Tahun 2013, (2014). https://www.kemkes.go.id/resources/download/pusdatin/profil-kesehatan-indonesia/profil-kesehatan-indonesia-2013.pdf

[3] G.E. Sugirtha, G. Murugesan, Detection of tuberculosis bacilli from microscopic sputum smear images, in: 2017 Third International Conference on Biosignals, Images and Instrumentation (ICBSII), IEEE, CHENNAI, India, 2017: pp. 1–6.

[4] E. Priya, S. Srinivasan, Automated object and image level classification of TB images using support vector neural network classifier, Biocybern. Biomed. Eng. 36 (2016), 670–678.

[5] A. Rachmad, N. Chamidah, R. Rulaningtyas, Image Enhancement Sputum Containing Mycobacterium Tuberculosis Using A Spatial Domain Filter, IOP Conf. Ser.: Mater. Sci. Eng. 546 (2019), 052061.

[6] K.S. Mithra, W.R.S. Emmanuel, FHDT: Fuzzy and Hyco-entropy-based Decision Tree Classifier for Tuberculosis Diagnosis from Sputum Images, Sādhanā. 43 (2018), 125.

[7] E. Purwanti, W. Prihartini, Using learning vector quantization method for automated identification of mycobacterium tuberculosis, Indones. J. Trop. Infect. Dis. 3 (2015), 26-29.

[8] F. Arisgraha, W. Prihartini, A. Retna, Digital Detection System Design of Mycobacterium Tuberculosis Through Extraction of Sputum Image Using Neural Network Method. Indones. J. Trop. Infect. Dis. 3 (2015), 35-38.

[9] J.S.R. Dinesh, Rajesh K. B., Tuberculosis (TB) detection system using deep neural networks, Neural Comput. Appl. 31 (2019), 1533–1545.

[10] J. Ker, L. Wang, J. Rao, T. Lim, Deep Learning Applications in Medical Image Analysis, IEEE Access. 6 (2018), 9375–9389.

[11] K. Munir, H. Elahi, A. Ayub, F. Frezza, A. Rizzi, Cancer Diagnosis Using Deep Learning: A Bibliographic Review, Cancers. 11 (2019), 1235.

[12] K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778.

[13] M.I. Shah, S. Mishra, V.K. Yadav, A. Chauhan, M. Sarkar, S.K. Sharma, C. Rout, Ziehl–Neelsen sputum smear microscopy image database: a resource to facilitate automated bacilli detection for tuberculosis diagnosis, J. Med. Imag. 4 (2017) 027503.

[14] Shorten C., and Taghi M. K., A survey on image data augmentation for deep learning, Journal of Big Data. 1 (2019), 1-48.

[15] M.D. Zeiler, R. Fergus, Visualizing and Understanding Convolutional Networks, in: D. Fleet, T. Pajdla, B. Schiele, T. Tuytelaars (Eds.), Computer Vision – ECCV 2014, Springer International Publishing, Cham, 2014: pp. 818–833.

[16] F. Damayanti, A. Rachmad, Recognizing Gender Through Facial Image Using Support Vector Machine. J. Theore. Appl. Inform. Technol. 88 (2016), 607 – 612.

[17] W. Hsu, T. Yu, E-mail spam filtering based on support vector machines with taguchi method for parameter selection. J. Convergence Inform. Technol. 5 (2010), 78-88.

[18] A. Rachmad, N. Chamidah, R. Rulaningtyas, Mycobacterium Tuberculosis Identification Based on Colour Feature Extraction Using Expert System, Ann. Biol. 36 (2020), 196-202.