17

Andini_2.pdf

by

Submission date: 25-Mar-2021 08:12AM (UTC+0700)

Submission ID: 1541616809

File name: Andini_2.pdf (2.66M)

Word count: 3046

Character count: 16373

Backpropagation Performance Against Support Vector Machine in Detecting Tuberculosis Based on Lung X-Ray Image

1st Arizal
Electrical Engineering
Universitas Fajar
Makassar, Indonesia
arizal@unifa.ac.id

2nd Andani Achmad
Electrical Engineering
Universitas Fajar
Makassar, Indonesia
andani@unhas.ac.id

3rd Andini Dani Achmad
Electrical Engineering
Universitas Fajar
Makassar, Indonesia
andini.achmad@unhas.ac.id

***Abstract*—Tuberculosis (TB) is known as an infectious disease caused by bacterium Mycobacterium Tuberculosis. It is one of the highest diseases that occur in Indonesia. The lung disease can be identified by analyzing the x-ray image of the lung.The problem that followed is that the x-ray imageswere analyzed separately by the specialist physician at separate times, so the patient should consult a doctor after getting the x-ray image. In this study, we create a modeling design that can detect TB disease early by using artificial neural network method that is backpropagation by using *Matlab Software*, furthermore analyze the performance of the modeling based on the level of accuracy.In training process this system uses 441 images while for the test used 221 x-ray images.The system's phases were started with preprocessing including median filter process and histogramequalization to improve image quality. The results of preprocessing is then classified with Backpropagation algorithm through training process. The results showed that TBC detection system can be built using backpropagation method with 4400 hidden layer hidden neurons with accuracy of 81.45% from the test process result. The accuracy of NN Backpropagation is better than SVM method whose accuracy reachesof 78.73%.**

Keywords—artificial neural network,backpropagation,accuracy,detection,tuberculosis

I. Introduction

In medicine, tuberculosis (TB) is known as an infectious disease caused by the microbacterial tuberculosis bacteria. This bacterium is a very strong bacillic bacteria that requires a long time to treat it, so early detection is very important for early prevention.

Abnormal of Tissue in the lungs can be detected early through X-rays. To obtain x-ray images was carried out by hospital radiology to the decision-making results of the diagnosis were carried out by expert doctors based on abnormalities of lung tissue by looking at the characteristics in the x-ray image.

Typical sign of pulmonary tuberculosis in an x-ray image as shown in circular white spots like stains on clothing. Therefore, the quality of the image produced by the X-ray machine plays a crucial role in the diagnosis process. The presence of noise in the image and lighting that is not good, it will affect the diagnosis of a specialist. Another checkup that supports the diagnosis of tuberculosis is by examining BTA sputum by conducting a laboratory test by examining the cough phlegm of the patientThe problem that occurs later is that for x-ray images issued by hospital radiology, the diagnosis process is generally carried out separately by the doctor at different times, so the patient must consult a specialist after obtaining an x-ray image. The manual detection process also causes differences in diagnosis between doctors. So the digital medical image processing can help improve the quality of reading x-ray images and can be considered or a second opinion for doctors.

In this study, TB detection was carried out using the median filter method and equalization histogram for preprocessing processes, which then the output from the preprocessing process was classified using the backpropagation neural network method, which will then detect images classified as images with suspect TB or normal as a result of early detection. Early detection results are included in the Early Detection Results Information System Provider of Tuberculosis that can be accessed by doctors and patients. This study also shows a comparison of the accuracy of the backpropagation neural network method with SVM in detecting TB.

II. Literature Review

A. *Tuberculosis (TBC)*

According to Misnadiarly[1], TB disease (Tuberculosis) is a chronic disease (chronic) that has long been known by the wider community and feared because it is contagious. In 1882, Robert Koch had conclusively provided evidence that tuberculosis was an infectious disease caused by bacteria called Mycobacterium tuberculosis. The bacterium is red, rod-shaped and acid resistant called Acid Resistant Basil (BTA).

The general symptoms of Tuberculosis are as follows

[2].

1) Coughing phlegm, the first complaints of coughing and the most common complaints. If TBC is severe enough, coughing will be accompanied by blood.
2) Shortness of breath, this complaint is found when pulmonary parenchymal damage is widespread or because there are other things that accompany it, such as pleural effusion (presence of fluid in the chest), pneumothorax (airway filled pleural cavity), anemia, and others.
3) chest pain in tuberculosis including mild pleuritic pain (chest infected with bacteria). These symptoms arise when the respiratory system in the pleura is exposed to tuberculosis.
4) Fever in the afternoon or evening, a complaint that is often encountered and usually arises in the afternoon or evening similar to influenza fever, loss of arising, and the longer the attacks are longer, while the free period of attacks is shorter.
5) Other complaints that usually arise are night sweats, anorexia, weight loss, and weak bodies (malaise). The onset of complaints is usually gradual appearing in a few weeks-months

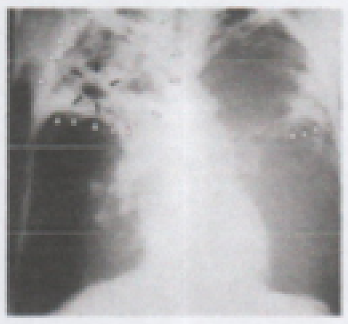

Figure 1. The X-ray results of Tuberculosis sufferer. (Sumber :Naing, W. &Htike, Zaw.[7])

White arrows indicate infection in both lungs. A black arrow indicates a hole that has been formed.

B. *Image Processing*

Image Processing is the process of improving image quality so that it is easily interpreted by humans or computers. Image processing techniques by transmitting images into other images, for example: image compression (image compression). Image processing is the initial process (preprocessing) of computer vision.

Furthermore, the results of image processing will be continued in the pattern recognition process is the next stage or analysis of image processing. Grouping of numerical and symbolic data (including images) automatically by the computer so that an object in the image can be recognized and interpreted [1].

Basics of Image Processing

To get results that are in accordance with the desired goals, the images stored in the form of digital files are generally processed first with the aim of improving the quality of the image.

Image Restoration

Image restoration is an important operation in image processing. The operation is used to reduce noise at the time of data acquisition, often when data acquisition is done the resulting image does not match what was previously desired. There are several causes of discrepancies such as a camera lens that is less clean so that spots appear (noise) and there is interference when sending data.

Image Segmentation

Segmentation is a technique for dividing an image into several regions where each region has similar attributes. In this section a number of segmentation techniques are covered, including floating, connected component tagging, cluster based segmentation and Hough transformation.

The thresholding process will produce a binary image, that is, an image that has two gray level values, black and white. In general, the process of grayscale imagery to produce binary images as follows:

$$g(x,y) = \{1 \text{ if } f(x,y) \geq T\}$$
$$g(x,y) = \{0 \text{ if } f(x,y) < T\}$$

where g(x,y) is a binary image from the grayscale image f (x, y), and T represents the threshold value. The value of T plays a very important role in the mining process. The quality of binary images is very dependent on the value of T used.

Artificial Neural Networks

Artificial neural network (ANN) is one part of artificial intelligence (artificial intelligence). ANN is defined as an information processing system that has characteristics resembling human neural networks and represents the human brain which always tries to simulate the learning process in the human brain. This artificial term is implemented using a computer program that is able to complete a number of calculation processes during the learning process.

AlgoritmaBackpropagation

Backpropagation is one of the training methods on neural networks, where the characteristics of this method are minimizing errors in the output generated by the network .Backpropagation trains the network to get a balance between the ability of the network to recognize the patterns used during training and the ability of the network to respond correctly to input patterns that are similar (but not the same) to the patterns used during

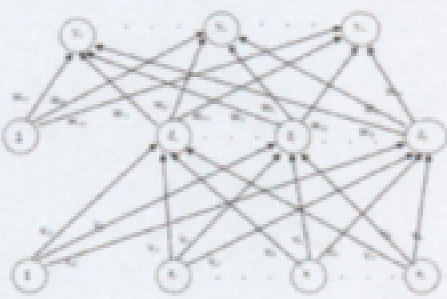

Figure 2. Arcitectureof Backpropogation

vji is the line weight of the input unit *xi* to the hidden unit layer *zj* (*vj0* is the line weight that connects the bias in the input unit to the hidden unit layer *zj*). *wkj* is the weight of the hidden layer *zj* to the output unit *yk* (*wk0* is the weight of the bias in the hidden layer to the output unit *zk*).

The backpropagation algorithm uses an output error to change the value of its weights in the backward direction. To get this error the forward propagation stage must be done.

III. Research Methods

A. *Stage of Research*

Thisresearchwillbeconductedbydevelopingfromseveralpr eviousstudies, in this study usingthebackpropagation neural networkmethodwiththestagescarriedout in this study as follows:

1. Literature Study
2. Data Collection
3. System Analysisand Design
4. System Testing
5. Interface Design
6. Report

B. *Perancangan Sistem*

The system used consists of a training process and a test process. Both the training process and the test begin with initial processing or also called pre-processing before being classified using the backpropagation JST algorithm. Pre-processing is done to improve image quality and take characteristics or characteristics of a digital image. The process that occurs in preprocessing is grayscale, median filter, equalization histogram, image conversion to black and white, and resize. The results of preprocessing then become input to the JST feature classification process with the backpropagation algorithm. In the training process, the configuration of backpropagation neuron architecture is carried out to obtain maximum accuracy.

IV. Results and Discussion

A. *Image Processing Results*

a. *Grayscale results*

The earliest stages in image processing are the process of converting images to grayscale images. Only part of the *x-ray* image is used in this study which has the type of grayscale image so that it is necessary to do a grayscale process for other images. Grayscale image is a digital image that has only one channel value in each pixel, meaning that the value of Red = Green = Blue. These values are used to indicate the intensity of the color.

To change the color image that has a matrix value of each R, G, and B to gray scale with a value of s, then the conversion can be done by taking the average values of r, g, and b so that they can be written as :

$$I(i,j) = \frac{R(i,j) + G(i,j) + B(i,j)}{3}$$

I (i.,j) = The intensity value of the grayscale image

R (i,j) = the value of red intensity from the original image

G (i,j) = green intensity value from the original image

B (i,j) = blue intensity value from the original image

B. *Results of the Classification of the Backpropagation JST Program and Comparison with the SVM Method*

The design of backpropagationb JST program consists of the process of training (training) and testing (testing). The data used in this study are divided into training data and test data. The training data used in the training process were 441 x-ray images consisting of 210 normal images and 231 abnormal images. The test data in the testing process were 221 x-ray images consisting of 116 normal images and 105 abnormal images. The backpropagation JST program is trained by using 100 to 5000 hidden neurons.

In the training process in this program, the input values are taken from the pixel values of the results of the black and white process that are converted into vector shapes. This feature vector will then experience network training repeatedly until the number of iterations (epoch) has been completed or the error value has reached the given tolerance. If the results of pattern recognition in this process have reached 100 percent; i.e. when the error is smaller than tolerance, the weights and biases can be stored and then used in the testing process.

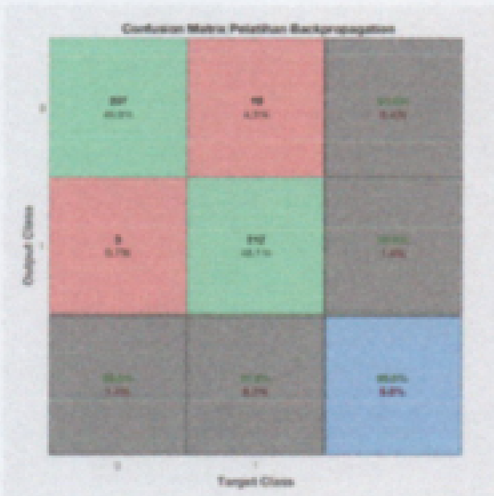

Figure 3. Confusion ofJST Backpropagation

Figure 3 shows Confusion Matrix classification is shown for training data using the Backpropagation JST method. In Confusion Matrix the line represents the predicted class (output class) while the column represents the actual class (target class). Green diagonal cells represent the correct classification results while pink diagonal cells represent the wrong classification results. In Confusion Matrix above 0 represents a normal image and 1 represents an abnormal image. In Figure 19 shows the system produces True

Positive as many as 207 images and True Negative as many as 212 images. So by using a formula.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
$$= \frac{207 + 212}{207 + 212 + 3 + 19} = 0{,}9501$$
$$= 95{,}01\%$$

It can be seen that the overall accuracy of the system in the training process classifies correctly between normal and abnormal images is 95.01%. The accuracy of the system is described with the following sensitivity and specificity values.

$$Sensitivitas = \frac{TP}{TP + FN} = \frac{207}{207 + 19} = 91{,}5\%$$

$$Spesifitas = \frac{TN}{FP + TN} = \frac{212}{3 + 212} = 98{,}60\%$$

The accuracy of the system is described as the system successfully classifying 98.6% of 210 normal data as normal data and 91.5% of 231 abnormal data as abnormal data. This value is obtained by using neurons as many as 4400 which are then applied to the testing process.

In the testing process, the backpropagation JST program identifies x-ray images that are not used in the training process. Program performance is calculated using the value of accuracy and error. Where the level of accuracy is the percentage of x-ray images that are successfully diagnosed correctly by the backpropagation JST program. The research that has been done on this artificial neural network system is to change the number of neurons in hidden layers (hidden layer).

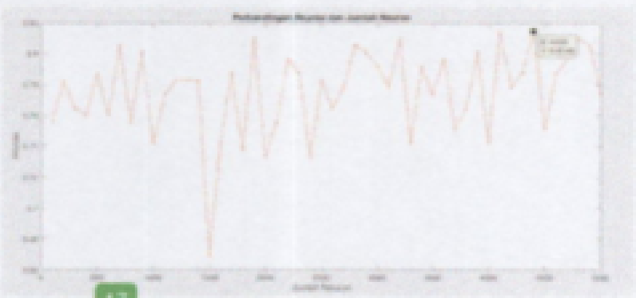

Figure 4. Comparison of Accuracy with the Number of Hidden Neurons

Figure 4 shows a comparison of accuracy with the number of hidden neurons used in the testing process for backpropagation JST programs. From Figure 20 above it is obtained that in the scale of hidden neurons from 0 to 5,000 the accuracy that occurs is in the interval of 66.97% to 81.45%. Where when the lowest value is obtained when the number of hidden neurons is 1500 with an accuracy of 66.97%. The highest accuracy value is reached by the number of hidden neurons as many as 4400 neurons, namely 81.45%.

Figure 5. *Confusion Matrix*of JST *Backpropagation*

Figure 5 shows Confusion Matrix classification for test data using the Backpropagation JST method. In Confusion Matrix the line represents the predicted class (output class) while the column represents the actual class (target class). Green diagonal cells represent the correct classification results while pink diagonal cells represent the wrong classification results. In Confusion Matrix above 0 represents a normal image and 1 represents an abnormal image. In Figure 24 shows the system produces True Positive as many as 97 images and True Negative as many as 83 images. So by using a formula.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
$$= \frac{97 + 83}{97 + 83 + 19 + 22} = 0.8145$$
$$= 81.45\%$$

It can be seen that the overall accuracy of the system in classifying correctly between normal and abnormal images is 81.45%. The accuracy of the system is described with the following sensitivity and specificity values.

$$Sensitivitas = \frac{TP}{TP + FN} = \frac{97}{97 + 19} = 83.62\%$$

$$Spesifitas = \frac{TN}{FP + TN} = \frac{83}{83 + 22} = 79.05\%$$

The accuracy of the system is described as the system successfully classifying 83.62% of 116 normal data as normal data and 79.05% of 105 abnormal data as abnormal data. The Backpropagation ANN program is then compared with the Polynomial Support Vector Machine (SVM) method for performance validation. The SVM method used is trained using the same data in the Backpropagation ANN training process. The highest accuracy using the SVM method is 78.7% by dividing training data and the same test with JST Backpropagation.

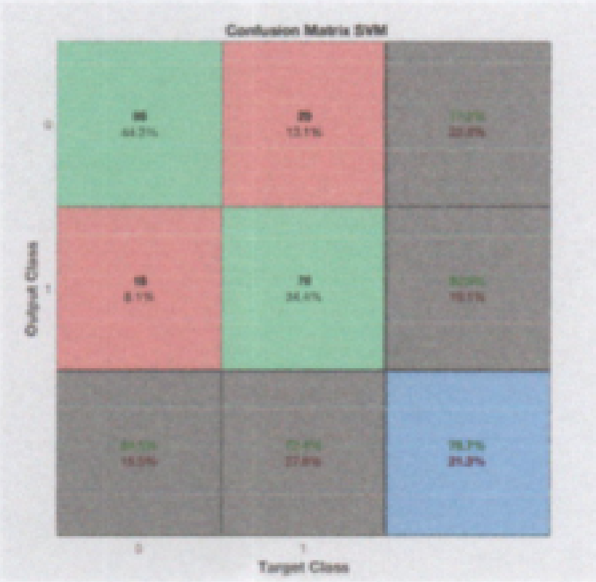

Figure 6. *Confusion Matrix* ofSVM Program

In the figure 6, the system produces 98 positive images and 76 negative images. So by using a formula.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
$$= \frac{98 + 76}{98 + 76 + 29 + 18} = 0.7873$$
$$= 78.73\%$$

9

It can be seen that the overall accuracy of the system in classifying correctly between normal and abnormal images is 78.73%. The accuracy of the system is described with the following sensitivity and specificity values.

$$Sensitivitas = \frac{TP}{TP + FN} = \frac{98}{98 + 18} = 84{,}5\%$$

$$Spesifitas = \frac{TN}{FP + TN} = \frac{76}{29 + 76} = 72{,}4\%$$

Conclusion

The accuracy of the system described as a system managed to classify 84.5% of 116 normal data as normal data and 72.4% of 105 abnormal data as abnormal data.

TB disease detection system from lung x-ray images can be built by combining several image processing methods as stages of image quality improvement and JST backpropagation method as image classification stages using Matlab software.

The TB disease detection system 21 is built using the JST backpropagation method with the number of 4400 hidden layer neurons having an accuracy rate of 81.45%. The accuracy of the JST backpropagation method is better than the detection system using the SVM method whose accuracy reaches 78.73%.

The TB detection system built using a friendly user interface makes it easier for users to use a TB detection program based on lung x-ray images.

References

[1] K, Abdul & S, Adhi. 2013. Teori dan Aplikasi Pengolahan Citra. Yogyakarta: Penerbit Andi Yogyakarta

[2] Abidin, Z. 2010. Pengenalan Ekspresi Wajah Menggunakan Metode Fisherface dengan Pendekatan Jaringan Syaraf Tiruan Back Propagation. Tesis. Yogyakarta: FMIPA Universitas Gajah Mada.

[3] Amin, S. 2012. Sistem Deteksi Dini Hama Wereng Batang Coklat Menggunakan Jaringan Syaraf Tiruan Backpropagation. Skripsi. Semarang: FMIPA Universitas Negeri Semarang.K. Elissa, "Title of paper if known," unpublished.

[4] Hartono, A. F. 2012. Implementasi Jaringan Syaraf Tiruan Backpropagation sebagai Sistem Pengenalan Citra Daging Sebagai Upaya Mengenali Daging Sapi Asli dan Daging Sapi Palsu di Pasar Tradisional Kota Semarang. Skripsi. Semarang: FMIPA Universitas Negeri Semarang..

[5] Mardianto, Is & Pratiwi, D. 2008. Sistem DeteksiPenyakit Pengeroposan Tulang dengan Metode Jaringan Syaraf Tiruan Backpropagation dan Representasi Ciri dalam Ruang Eigen. CommIT, Vol.2 No.1 Mei 2008, hlm 69-80.

[6] Arisgraha, F., Widiyanti, P., & Apsari, R. 2012. Digital Detection System Design of Mycobacterium Tuberculosis Through Extraction of Sputum Image Using Neural Network Method. Indonesian Journal of Tropical and Infections Disease, Vol. 3 No. 1 January-March 2012.

[7] Naing, W. & Htike, Zaw. 2014. Advances in Automatic Tuberculosis Detection in Chest X-Ray Images. Signal & Image Processing: An International Journal (SIPIJ) Vol 5, No. 6, December 2014.

[8] Achmad, A., Warni, E., & friends. 2013. Deteksi Karies Gigi Menggunakan Jaringan Syaraf Tiruan Dengan Metode Backpropagation.

Andini_2.pdf

ORIGINALITY REPORT

20%	%	20%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Jimmy Tandean, Endry, Alvin Wijaya, Winata Gunawan, Mawaddah Harahap. "Vehicle Collision Detection Application Through Collision Video Files with Quadtree Algorithms", Journal of Physics: Conference Series, 2019 Publication	3%
2	Arif Fridasari, Ema Utami. "Recognition Pattern of Arca Siwa Prambanan Temple with Canny and Backpropagation Algorithm", 2019 International Conference on Information and Communications Technology (ICOIACT), 2019 Publication	2%
3	I Wayan Agus Surya Darma, Ni Putu Sutramiani. "Segmentation of Balinese Script on Lontar Manuscripts using Projection Profile", 2019 5th International Conference on New Media Studies (CONMEDIA), 2019 Publication	1%
4	A P Widowo, E A Sarwoko, Suhartono. "Performance comparison of machine learning methods for prediction of estimating water	1%

production", Journal of Physics: Conference Series, 2019
Publication

5 B. M. Pratama, A. Damayanti, E. Winarko. "Signature image identification using hybrid backpropagation with firefly algorithm and simulated annealing", AIP Publishing, 2021 — 1%
Publication

6 D R S Saputro, I M Putri, Sutanto, N H Noor, P Widyaningsih. "Generalized space time autoregressive (gstar)-artificial neural network (ann) model with multilayer feedforward networks architecture", IOP Conference Series: Earth and Environmental Science, 2019 — 1%
Publication

7 Christin Panjaitan, Aldo Silaban, Mikhael Napitupulu, Joni Welman Simatupang. "Comparison K-Nearest Neighbors (K-NN) and Artificial Neural Network (ANN) in Real Time Entrants Recognition", 2018 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), 2018 — 1%
Publication

8 Wina Safutri, Nova Kurniati, Fransisca Srioetami Tanoehardjo, Kemas Yakub Rahadiyanto. "IFN-gamma and IL-4 Secretion After Stimulation of — 1%

EC610 Fusion Antigens (ESAT-6-CFP-10) in Patients With Active Pulmonary TB And Latent TB", Biomedical Journal of Indonesia, 2020
Publication

9 "Advances in Visual Computing", Springer Science and Business Media LLC, 2020 1%
Publication

10 Prashant Shukla, Abhishek Verma, Abhishek, Shekhar Verma, Manish Kumar. "Interpreting SVM for medical images using Quadtree", Multimedia Tools and Applications, 2020 1%
Publication

11 Bhaskar Dey, Malay Kumar Kundu. "Turning video into traffic data – an application to urban intersection analysis using transfer learning", IET Image Processing, 2019 1%
Publication

12 Ceri Ahendyarti, Romi Wiryadinata, Neneng Rohana, Fadil Muhammad. "Cloud Classification from NOAA Satellite Image Using Learning Vector Quantization Method", 2020 2nd International Conference on Industrial Electrical and Electronics (ICIEE), 2020 1%
Publication

13 Anita Sindar R M Sinaga. "The Comparison of Signature Verification Result Using 2DPCA Method and SSE Method", International Journal 1%

of Artificial Intelligence Research, 2018
Publication

14	S Purwiyanti, RP Farisi, F A Setyawan, S R Sulistyanti, Y Yuniati, U Murdika. "The Use of Background Subtraction Method in Model Design of Traffic Light Timer Control System", Journal of Physics: Conference Series, 2019 Publication	1%
15	S.M. Swain, J. Baselga, D. Miles, Y.-H. Im, C. Quah, L.F. Lee, J. Cortés. "Incidence of central nervous system metastases in patients with HER2-positive metastatic breast cancer treated with pertuzumab, trastuzumab, and docetaxel: results from the randomized phase III study CLEOPATRA", Annals of Oncology, 2014 Publication	1%
16	A D Achmad, A Achmad, Z M Putra. "Interactive Learning Media Based on Android for Computer Programming Course", IOP Conference Series: Materials Science and Engineering, 2020 Publication	1%
17	Cheonghwan Hur, Bunyodbek Ibrokhimov, Sanggil Kang. "N3-CPL: Neuroplasticity-based neuromorphic network cell proliferation learning", Neurocomputing, 2020 Publication	1%
18	Fei Gao, Zheng Li, Gang Xiao, Xiaoyang Yuan,	

Zhenggao Han. "An online inspection system of surface defects for copper strip based on computer vision", 2012 5th International Congress on Image and Signal Processing, 2012
Publication

<1%

19 Yinxue Zhang. "Image Restoration Based on Robust Error Function and Particle Swarm Optimization-BP Neural Network", 2008 Fourth International Conference on Natural Computation, 10/2008
Publication

<1%

20 Abdul Fadlil, Rusydi Umar, Sapriani Gustina. "Mushroom Images Identification Using Orde 1 Statistics Feature Extraction with Artificial Neural Network Classification Technique", Journal of Physics: Conference Series, 2019
Publication

<1%

21 Shoffan Saifullah, Vynska Amalia Permadi. "Comparison of Egg Fertility Identification based on GLCM Feature Extraction using Backpropagation and K-means Clustering Algorithms", 2019 5th International Conference on Science in Information Technology (ICSITech), 2019
Publication

<1%

22 Victor Marudut Mulia Siregar, Heru Sugara.

<1%

"Implementation of artificial neural network to assesment the lecturer's performance", IOP Conference Series: Materials Science and Engineering, 2018
Publication

---

23 Oskar Hutagaluh, Sardjana Orba Manullang, Muliyadi Hamid, Abdul Samad Arief, Dayan Abdurrahman. "HOW THE RESEARCH SKILLS DEVELOPMENT FRAMEWORK MAY BE USEFUL FOR INDONESIAN HIGHER EDUCATION LEARNING PRACTICE", Nidhomul Haq : Jurnal Manajemen Pendidikan Islam, 2020
Publication
<1%

Exclude quotes	On	Exclude matches	< 5 words
Exclude bibliography	On		