International Journal of Industrial Optimization
Volume 2 Issue 2 September 2021 pp. 141-150

P-ISSN 2714-6006
E-ISSN 2723-3022

# The hybrid design of supervised learning algorithm for design and development in classifications a defect in clay tiles

Murman Dwi Prasetio*, Rais Yufli Xavierullah, Haris Rachmat, Wiyono, Denny Sukma Eka Atmaja

School of Industrial and System Engineering, Telkom University, Bandung, Indonesia

murmandwi@telkomuniversity.ac.id; raisxavier12@gmail.com; harisrachmat@telkomuniversity.ac.id; wiyono@telkomuniversity.ac.id; dennysukma@telkomuniversity.ac.id;

*Corresponding Authors: murmandwi@telkomunviersity.ac.id

ARTICLE INFO

**Keywords**
Clay Tile;
Support Vector Machine;
Local Binary Pattern;
Supervised Learning.

**Article history**
Received:
July 14, 2021
Revised:
July 18, 2021
Accepted:
August 21, 2021
Available online:
August 31, 2021

ABSTRACT

The strength of the company's competitiveness is needed because the current industrial development is very rapid. This is necessary with the aim of maintaining the quality and quantity of the products produced according to company standards. One of the companies that must maintain the quality and quantity is PT. XYZ is a clay tile company. The classification of products used by this company to maintain good quality is three classes, namely good tile, white stone tile, cracked tile. However, the practice of quality control based on classification still uses the traditional way by relying on sight. It can increase errors and slow down the process. This can be overcome with artificial visual detectors. This is a result of the rapid development of automation. So to detect defects can use image preprocessing, supervised learning algorithms, and measurement methods. Support Vector Machine (SVM) is used in this study to perform classification while feature extraction on clay tiles used the Local Binary Pattern (LBP) method. The algorithm is made using python while for image retrieval the raspberry pi is used. The linear kernel on the SVM algorithm is used in this study. The conclusion in this study obtained 86.95% the highest accuracy with a linear kernel. It takes 10.625 seconds to classify.

## 1. Introduction

The strength of the company's competitiveness is needed because the current industrial development is very rapid. This is necessary with the aim of maintaining the quality and quantity of the products produced according to company standards. In keeping with this, a change was made from production activities that initially used human (manual) labor to become automatic (Bilghifary et al., 2015). One of the companies that must maintain the quality and quantity is PT. XYZ is a clay tile company. Critical companies in Indonesia are non-metal goods companies in micro and small industries (BPS, 2019). There is an increase of 10.12% from 2015 to 2019. If

P-ISSN 2714-6006
E-ISSN 2723-3022

this happens then production will increase. But the problem is when there is no demand from consumers, it will cause a company to experience losses (Junaidi et al., 2019)[9]. One of the keys to getting consumers is maintaining quality (Kusumawati & Fitriyeni, 2017).

In maintaining the quality of a company can use quality control which is expected to prevent defects in products that can cause losses in terms of material and labor. PT. XYZ itself has classifications for products, namely good tiles, white stone tiles, and cracked tiles according to Table 1 which will become a reference for the quality of tiles. However, the practice of quality control based on classification still uses the traditional way by relying on sight. The use of sight vision can have risks such as increased operating costs due to incorrect checks, failure to get business and rework (Prasetio & Xavierullah, 2020). It can increase errors and slow down the process (Ragab & Alsharay, 2017). Fatigue and frustration when carrying out detailed inspections using vision can increase errors (Oktaviani & Budiman, 2019)[13]. This can be overcome with artificial visual detectors. This is a result of the rapid development of automation. So to detect defects can use image preprocessing, supervised learning algorithms, and measurement methods (Li et al., 2018).

Table 1. Type of Clay

| Number | Type of Clay | Figure |
|---|---|---|
| 1 | Good |  |
| 2 | Whitestone |  |
| 3 | Crack |  |

The creation of models for classifying an object in the same class based on the characteristics is called classification (Aprilla et al., 2018). SVM is a method that can be used for classification (Aprilla et al., 2018). SVM is used because it can always achieve the same solution on each trial. This is due to the ability of SVM to find optimal global properties (Shadika, 2017). Therefore, to classify clay tiles; good tiles, tiles with white stone defects, tiles with crack defects in this study used SVM intending to increase accuracy.

## 2. Literature Review

Research that conducted by Jiang & Yin (2018) to design classification of defective tiles in sleet. This research was conducted by introducing a defect detector using the Bayesian method. To improve its accuracy This study uses deletion of pixels that exist in images that are not a

P-ISSN 2714-6006
E-ISSN 2723-3022

handicap. This study gets an accuracy of value in its own research proposed tool is successful in identifying outline. Research conducted by Oktaviani & Budiman (2019) designing a system that can classify types of sandstone. Aim of this research as a tool for geologists to identify species of rock. The method used for feature extraction is Local Binary Pattern (LBP) while the method used in classification is method Support Vector Machine (SVM). This research gets the highest accuracy amounted to 93.40% with a computing time of 0.98 seconds. A research was conducted by Luo (2014) to detect ceramic tubes using the SVM-QPSO hybrid algorithm. This study uses image data totaling 500 ceramic images tubes with 5 types of ceramic tubes are used. The classification of this research gets an accuracy value of 85.8%.

As for research conducted by Bong et al. (2019) by making tools sight-based automation in detecting and classifying defects found on fabric. The method used in this research is Support Vector Machine (SVM). Data used in this study were 2500 data with training data 2000 data and test data 500 data with each defect as many as 5 classifications of defects. This study gets an accuracy value of 98%. Research conducted by Prasetio, (2020) approach machine learning used tile inspection case study. The method used in this research is Local Binary Pattern (LBP) and Support Vector Machine (SVM). This study used an image with three types of tiles. Result classification in this research gets an accuracy of 76.67%. Furthermore, research was conducted by Jawahar et al. (2015) with classifying fabrics using wavelet feature extraction method and classification using Support Vector Machine (SVM) method. Data used in this study were 500 and 200 defects data is not defective with image resolution used is 256 x256 pixels. In its classification, there is the highest accuracy value of 98.84%. In addition, there is research conducted by Ragab & Alsharay (2017) by developing algorithms to detect and classify defects contained on ceramics using Compute Unified Device Architecture (CUDA). The defects studied were crack defects and bitnik defects freckles. The highest accuracy value was obtained at 73%. The research this time will use three kernels as research material, namely linear, polynomial, RBF with the following formula Helyudanto et al. (2019):

1. *Linear*:
$$K(x, y) = xy \tag{1}$$
2. *Polynomial*:
$$K(x, y) = (x.y + c)^d \tag{2}$$
3. Rbf:
$$K(x, y) = \exp(\gamma \|x - y\|^2) \tag{3}$$

Information:
*K* = Kernel
*x,y* = Dot product
c = Konstanta
d = Degree of kernel
γ = Gamma

## 3. Research Methodology

### *3.1 Object of Research*

The research object used was clay tiles. Tiles are those that are directly inspected by an operator who is still using eyesight. Types studied were good tiles, cracked deformed tiles, white stone deformed tiles. The research object can be seen in Figure 1.

P-ISSN 2714-6006
E-ISSN 2723-3022

(a) (b) (c)

Figure 1. The object of Research (a) Good, (b) Whitestone, (c) Crack

*3.2 System Design*

The design system used in this study is shown in Figure 2. In this research information topology, a laptop is used as an algorithm maker that is made and runs the system when the training and testing process is carried out using data that is already owned. Model results from the training process will be stored in cloud storage which will be used during the real-time experiment.

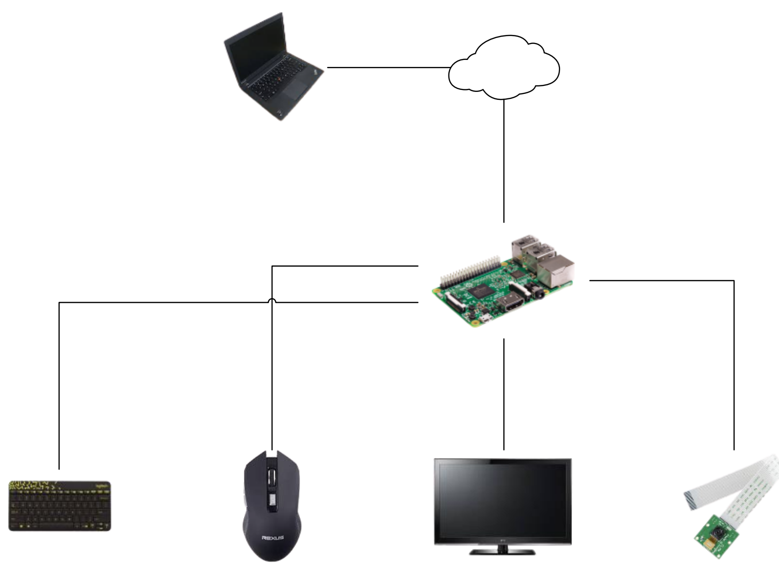

Figure 2. Information Topology

Data that has been stored in cloud storage will be retrieved by raspberries. Raspberries that are used cannot read "ssh" file so raspberries cannot use a laptop as controller of raspberry, so raspberries in this study must use or devices, namely monitor screen, keyboard, mouse. As an image taker, the Raspberry camera module is directly used. As support for the camera when taking images, a research rig is made as shown in Figure 3. The basis for selecting the rig is the research of Atmaja & Herliansyah (2015) with a height of 50 cm and the camera used has a resolution of 360 x 480 pixels.

P-ISSN 2714-6006
E-ISSN 2723-3022

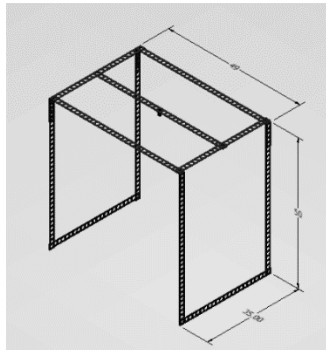

Figure 3. Research Rig

### *3.3 Classification Scenarios*

Figure 4 is a classification scenario image in determining where or not a tile is defective. The precarious clay tile image is taken in the first step because it is the input in this process. The next step is to convert the image to grayscale. So that after that, feature extraction can be carried out for reference in the classification process. Furthermore, the results of feature extraction are stored as a model to carry out direct tests. The last is to calculate the prediction results against the classification results.

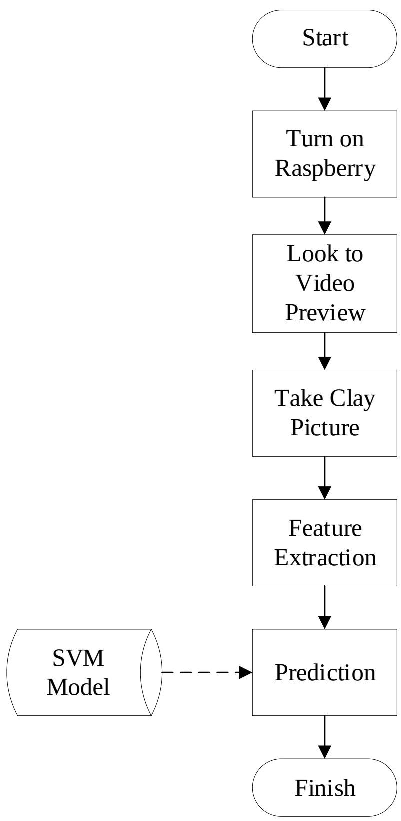


Figure 4. Classification Scenarios

*3.4 Dataset*

To improve accuracy results when classifying, data augmentation can be used by increasing the data used. Each image is flipped horizontally and vertically at random, while the data will be reproduced 3 times. The data is split into two, training and test data. Training data is data that is used to be trained with the SVM method. Test data is the data used to test the suitability of the label. Data splitting is done randomly by the system. Table 2 below is a split of test data and training data.

Table 2. Labeling and Splitting Data

| Class | Label | Training Data | Test Data |
|---|---|---|---|
| **Good** | 0 | 54 | 24 |
| **Whitestone** | 1 | 62 | 28 |
| **Crack** | 2 | 87 | 36 |
| **Total Data** | | 203 | 88 |

*3.5 Feature Extraction*

Feature extraction is done to simplify the classification of clay roof tiles. The Local Binary pattern in Figure 5 is the feature extraction used. Local Binary Pattern is done by determining the middle value of the grayscale image pixels. Make changes to the neighboring values by comparing the values in the middle. If the neighbor value is greater then it will be given a value of "1" otherwise it will be given a value of "0. Then the calculation is done to get the new pixel value by multiplying the changed binary value. This value is the new pixel value or the value from the LBP process.

Figure 5. Feature Extraction

P-ISSN 2714-6006
E-ISSN 2723-3022

## 4. Results and Discussion

### *4.1 Support Vector Machine*

This study uses a comparison between several kernels used, namely linear kernel, RBF kernel, and polynomial kernel. A confusion matrix is used to see system can classify well in predicting the accuracy of each class (Prasetio et al., 2018). The confusion matrix on the linear kernel can be seen in Table 3. 24 tiles that should be good class are predicted to be 16 good classes, 5 white stone defect classes, 3 cracked defect classes. 28 total tiles should be white stone class are predicted to be 0 good classes, 26 white stone defect classes, 2 cracked defect classes. 36 Clay tiles that should have a cracked defect class are predicted to be 0 good class, 1 white stone class, 35 cracked defects class.

Table 3. Confusion Matrix Kernel Linear

| Correct Classification | | Predicted Class | | |
|---|---|---|---|---|
| | | **Good** | **Whitestone** | **Crack** |
| **Actual Class** | **Good** | 16 | 5 | 3 |
| | **Whitestone** | 0 | 26 | 2 |
| | **Crack** | 0 | 1 | 35 |

The confusion matrix in the RBF kernel can be seen in Table 4. 24 total tiles that should be good class are predicted to be 9 good classes, 4 white stone defect classes, 11 cracked defect classes. 28 total tiles should be white stone class are predicted to be 0 good classes, 26 white stone defect classes, 2 cracked defect classes. Furthermore, 36 total tiles should have a crack defect class which is predicted to be 0 good class, 0 white stone defect class, 36 crack defect class.

Table 4. Confusion Matrix Kernel RBF

| Correct Classification | | Predicted Class | | |
|---|---|---|---|---|
| | | **Good** | **Whitestone** | **Crack** |
| **Actual Class** | **Good** | 9 | 4 | 11 |
| | **Whitestone** | 0 | 26 | 2 |
| | **Crack** | 0 | 0 | 36 |

The confusion matrix on the polynomial kernel can be seen in Table 5. 24 tiles that should be good classes are predicted to be 14 good classes, 6 white stone defect classes, 4 cracked defect classes. 28 total tiles should be white stone class are predicted to be 0 good classes, 26 white stone defect classes, 2 cracked defect classes. Furthermore, there are 36 total tiles, which should have a crack defect class which is predicted to be 0 good class, 1 white stone defect class, 35 crack defect class.

Table 5. Confusion Matrix Kernel Polynomial

| Correct Classification | | Predicted Class | | |
|---|---|---|---|---|
| | | Good | Whitestone | Crack |
| Actual Class | Good | 14 | 6 | 4 |
| | Whitestone | 0 | 26 | 2 |
| | Crack | 0 | 1 | 35 |

After using the confusion matrix, the process of calculating accuracy is carried out as in Table 6 showing that the level of accuracy in each kernel is different. The linear kernel has an accuracy of 87.5% while the RBF kernel has an accuracy rate of 80.6%, and the polynomial kernel has an accuracy rate of 85.2%. The greatest level of accuracy is a linear kernel with an accuracy rate of 87.5%.

Table 6. Accuracy of Each Kernel

| Kernel Name | Accuracy |
|---|---|
| **Linear** | 86.95% |
| **RBF** | 80.6% |
| **Polynomial** | 85.2% |

*4.2 Testing Time Results*

The testing time for the real-time process is shown in Table 7 below. Obtained 0.71 seconds of image capture, 9.51 seconds of feature extraction, and 0.41 seconds of image prediction.

Table 7 Processing Time

| Process | Processing Time (seconds) |
|---|---|
| **Image Capture** | 0,71 |
| **Feature Extraction** | 9,515 |
| **Image Prediction** | 0,40 |
| **Total** | 10,625 |

## 5. Conclusion

The results of the design for the classification of defective clay tiles obtained an accuracy of 86.95% and a real-time testing time of 10.625 seconds. The method used is LBP for feature extraction and SVM for classification.

Further research is suggested to use a better camera to get better image results, use the light from the lights that have been installed so that the image quality does not look dark which can result in a classification prediction model, and change the research not to use video previews which can take a long time when the setting is unstable.

## References

Aprilla, S., Furqon, M. T., & Fauzi, M. A. (2018). Klasifikasi Penyakit Skizofrenia dan Episode Depresi Pada Gangguan Kejiwaan Dengan Menggunakan Metode Support Vector Machine ( SVM ). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *2*(11), 5611–5618.

Atmaja, D. S. E., & Herliansyah, M. K. (2015). Optimasi Proses Pengukuran Dimensi Dan Defect Ubin Keramik Menggunakan Pengolahan Citra Digital Dan Full Factorial Design. *Jurnal Teknosains*, *4*(2), 179–191. https://doi.org/10.22146/teknosains.7972.

Bilghifary, M., Rachmat, H., & Sjafrizal, T. (2015). *Perancangan User Requirements Specification ( URS ) Sistem Otomasi Terintegrasi Pada Stasiun Exturning, Drilling-Chamfering, Dan Threading Di Pt . Abc Design of User Requirements Specification ( URS ) Integrated Automation System in Exturning, Drillin*. *2*(2), 3923–3957.

Bong, H. Q., Truong, Q. B., Nguyen, H. C., & Nguyen, M. T. (2019). Vision-based Inspection System for Leather Surface Defect Detection and Classification. *NICS 2018 - Proceedings of 2018 5th NAFOSTED Conference on Information and Computer Science*, 300–304. https://doi.org/10.1109/NICS.2018.8606836.

BPS. (2019). *Perkembangan Indeks Produksi Industri Triwulanan Industri Mikro dan Kecil 2017-2019*. Badan Pusat Statistik. https://www.bps.go.id/publication/2019/12/16/4bbaf229c500ad2439aa73f3/perkembangan-indeks-produksi-triwulanan-industri-mikro-dan-kecil-2017-2019.html.

Helyudanto, D., Nhita, F., & Rohmawati, A. A. (2019). *Prediksi Penyebaran Demam Berdarah di Kabupaten Bandung dengan Metode Hybrid Autoregressive Integrated Moving Average ( ARIMA ) dengan Support Vector Machine ( SVM )*.

Jawahar, M., Babu, N. K. C., & Vani, K. (2015). Leather texture classification using wavelet feature extraction technique. *2014 IEEE International Conference on Computational Intelligence and Computing Research, IEEE ICCIC 2014*, 6–9. https://doi.org/10.1109/ICCIC.2014.7238475.

Jiang, F., & Yin, G. (2018). *Bayesian Outdoor Defect Detection*. *14*(8), 1–9. http://arxiv.org/abs/1809.01000.

Junaidi, J., Koriatul, J., & Sutrisno. (2019). Model Aplikasi Purchasing System Untuk Monitoring Stok Dalam Mengurangi Tingkat. *Sensi*, *5*(1), 86–98. http://ejournal.raharja.ac.id/index.php/sensi/article/download/745/565.

Kusumawati, A., & Fitriyeni, L. (2017). Pengendalian Kualitas Proses Pengemasan Gula Dengan Pendekatan Six Sigma. *Jurnal Sistem Dan Manajemen Industri*, *1*(1), 43.

 P-ISSN 2714-6006
E-ISSN 2723-3022

https://doi.org/10.30656/jsmi.v1i1.173.

Li, Z., Zhang, J., Zhuang, T., & Wang, Q. (2018). Metal surface defect detection based on MATLAB. *Proceedings of 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference, IAEAC 2018*, IAEAC, 2365–2371. https://doi.org/10.1109/IAEAC.2018.8577540.

Luo, X. C. (2014). A hybrid SVM-QPSO model-based ceramic tube surface defect detection algorithm. *Proceedings - 2014 5th International Conference on Intelligent Systems Design and Engineering Applications, ISDEA 2014*, 28–31. https://doi.org/10.1109/ISDEA.2014.15.

Oktaviani, I., & Budiman. (2019). *Klasifikasi Jenis Batuan Pasir Sedimen Melalui Pengolahan Citra Digital Menggunakan Metode Local Binary Pattern ( LBP ) Dan Support Vector Machine ( SVM )*.

Prasetio, M. D., Hayashida, T., Nishizaki, I., & Sekizaki, S. (2018). Deep belief network optimization in speech recognition. *Proceedings - 2017 International Conference on Sustainable Information Engineering and Technology, SIET 2017*, *2018-Janua*, 138–143. https://doi.org/10.1109/SIET.2017.8304124.

Prasetio, M. D., & Xavierullah, R. Y. (2020). An Approaching Machine Learning Model: Tile Inspection Case Study. *International Journal of Innovation in Enterprise System*, *4*(01), 12–22. https://doi.org/10.25124/ijies.v4i01.44.

Ragab, K., & Alsharay, N. (2017). Developing Parallel Cracks and Spots Ceramic Defect Detection and Classification Algorithm Using CUDA. *Proceedings - 2017 IEEE 13th International Symposium on Autonomous Decentralized Systems, ISADS 2017*, 255–261. https://doi.org/10.1109/ISADS.2017.14.

Shadika. (2017). *Optimasi Klasifikasi Cacat Pada Kain Tenun Gorden Menggunakan Metode Image Processing Dan Metode Artificial Neural Network di Pt Buana Intan Gemilang*.