Building Crack Due To Lombok Earthquake Classification Based on GLCM Features and SVM Classifier

by Prodi Teknik Informatika Ft-unram

Submission date: 22-Feb-2022 11:30AM (UTC+0700)

Submission ID: 1768077243

File name: ICAMIMIA IEEE Published.pdf (5.12M)

Word count: 4098

Character count: 21893

Building Crack Due To Lombok Earthquake Classification Based on GLCM Features and SVM Classifier

1st I Gede Pasek Suta Wijaya

Informatics Engineering Dept., Engeneering Faculty
University of Mataram, Jl. Majapahit 62
Mataram, Lombok NTB, INDONESIA
gpsutawijaya@unram.ac.id

1 3rd Ida Bagus Ketut Widiartha
Informatics Engineering Dept., Engeneering Faculty
University of Mataram, Jl. Majapahit 62
Mataram, Lombok NTB, INDONESIA
widi@unram.ac.id

2nd Chaerus Sulton
Informatics Engineering Dept., Engeneering Faculty
University of Mataram, Jl. Majapahit 62
Mataram, Lombok NTB, INDONESIA
chaerussulton@gmail.com

14th Ni Nyoman Kencanawati

Civil Engineering Dept., Engeneering Faculty

University of Mataram, Jl. Majapahit 62

Mataram, Lombok NTB, INDONESIA

nkencanawati@unram.ac.id

Abstract—Cracks classification on buildings caused by natural disasters such as earthquakes can be done manually by analyzing walls, beams, columns, and floors based on visual inspection of cracks diameter, depth, and length. The manual assessment method requires experts in structural engineering who have enough knowledge and experience in building damage assessment. To facilitate and overcome these problems, a crack classification system is developed by using a digital image processing approach (pattern recognition) that can classify cracks into the mild, 19 derate, or severe categories using GLCM features and SVM classifier. Based on the experimental results that the proposed method has appropriately worked for classification of two crack classes (mild and severe) indicated by 94.44% of accuracy, 88.89% of precision, and 100.00% of recall. While for three crack classes (mild, moderate, and severe) obtained the accuracy 81,48%, recall 81,41% and precision 88,09%. Furthermore, the proposed system also shows robust performance against large variability of crack and non-crack images, and the SVM classifier outperforms over the statistical-based classifier (LDA and QDA).

Index Terms—Cracks, Image Analysis, GLCM, SVM, classification

I. INTRODUCTION

An earthquake is a serigiof vibrations with specific frequencies (seismic waves) that occur on the surface of earth due to the sudden release of energy due to a fault or an explosion. The seismic waves are energy propagations caused by disturbances in the earths crust. Therefore, the area affected by the earthquake will have various damage such as landslides, loss to infrastructure such as roads, bridges, houses, damage to other buildings. Additionally, an earthquake also triggers a tsunami which has massive damaged power.

Likewise, the North Lombok Regency, West Nusa Tenggara province, starting from July 29, 2018 to August 30, 2018 has been hit by 1,973 earthquakes [1]. The affected area has

various damage ranging from landslides, infrastructure, and other buildings. In terms of buildings (house, shop, office, etc.) themselves, earthquakes cause mild, moderate, and severe damage which can be assessed from the crack level. In fact, the cracks can also be caused by not only earthquakes but also by the age of buildings (old building usually becomes fragile and easily collapse by a small earthquake). Then, it is required an assessment of whether the building is still feasible or not based on its crack information.

An effort to collect building damage based on its level of cracks after an earthquake can be done by building assessment. The assessment is performed by observing and analyzing the cracks in the walls, columns, beams, and floors manually. Commonly, the damage level is determined based on crack's length, width, and area (quantitative metrics), which is categorized as minor/mild, moderate, and severe cracks [2]. The manual building assessment method has its drawbacks because it requires experienced building structure experts. However, the assessment process requires a long time and is expensive due to limited experts and large affected area.

To overcome such kind of problem, a crack classification system based on digital image processing (pattern recognition) approach is proposed which can categorize a crack into a mild, moderate or severe, using the GLCM features and SVM classifier. The proposed scheme can work automatically and potentially to be developed for an automatic non-destructive building assessment after an earthquake. Such kind of system is essential to support the government/community in earthquake management, given the position of Indonesia that is vulnerable to shocks.

978-1-7281-3090-3/19/\$31.00 ©2019 IEEE

II. RELATED WORKS

Researches related to the use of the Gray Level Co-Occurrence Matrix (GLCM) as a texture feature extraction method have been carried out by several researchers [3]–[6] including handwriting recognition [7], and quality of beef classification using ultrasound image data [8], with an accuracy in the range of 70% to 97.9%. The type of data, amount of data, pre-processing techniques, and the number of features used cause differences of accuracy. However, the achievement proves that the GLCM feature is auspicious as unique features of an object.

Crack detection and classification techniques with quantitative analysis on an infrastructure (the road, bridge, pavement, building, railway track, tunnel, ship, vehicle, and aircraft) is an essential process in finding crack level using various approaches [2], [9]–[11]. A great review and analysis of crack detection have been successfully examined on 20 image processing algorithms [3]. From this review, it was found that the GLCM with ANN (Artificial Neural Network) methods are also included in the work for handling the crack detection for thermography, visual color and grayscale images of concrete blocks. However, its performance lack [19] accuracy (ranging from 71 to 75.2%) [4], [5]. Additionally, a deep-learning based approach for road crack detection has provided remarkable achievement when compared with features obtained with existing hand-craft techniques [12].

Furthermore, utilization Support Vector Machine (SVM) [13], [14] as a classification method has also been successfully carried out, namely the classification handwriting [7] and Color grading of beef fat [14] which results in accuracy ranging from 87.5% to 97.9%. Similar to before, the difference in results is caused by the type of data, the number of features, and the pre-processing techniques. Additionally, A method based on SVM approach with Gabor wavelet features has been successfully developed for detecting defects on textured surfaces which show its effectiveness over the learning vector quantization (LVQ) [15].

Based on the previous research, the texture features extracted by using the GLCM accompanied by SVM classification method is proposed for the crack level classification of Lombok earthquake damage building. The propose of this classification system is to find the compact and powerful scheme for crack level classification of damage building due to the earthquake. The compact scheme must require less memory, which is potentially developed for the smartphone.

III. THE PROPOSED SCHEME

The proposed scheme has three main processes: preprocessing, training, and classification as presented in Fig. 1.

A. Preprocessing

Preprocessing phase itself consists of resizing, binarization, morphological filtering, and segmentation processes. Each training image and test image must pass preprocessing phase to eliminate and reduce the obstacles that can improve the

performance of the classification. The resize, binarization, morphological filtering, and segmentation algorithms commonly can be found in digital image processing [16].

- Resize is a process of reducing the size and cropping of images so that the data used is standardized and has the same size for all cracked and non-cracked images, which is 227 27 pixels.
- 2) Binarization is a transformation process of the image from RGB color space to black and white based on 127 thresholds. It means that a pixel value which is less than 127 is changed to 0, and the otherwise is changed to 1.
- Morphological filtering is the process of removing noise or specks to obtain a brighter and cleaner image. In this case, the filtering method utilized is median filtering.
- 4) Segmentation is just the localization of the cracks in the image. It means that only part of the crack of the input image is taken. In this case, the input image is the output image of the Morphological filtering process.

The input and output example of pre-processing for the crack classification is presented in Fig. 2.

B. Feature Extraction

Feature extraction, which is part of a pattern recognition technique, is intended to extract unique values of an object. In the case of crack classification, \mathbf{a} e features are extracted by the GLCM method. The GLCM is a matrix that represents the frequency of appearance of two-pixel pairs with a certain intensity in the distance (d) and direction orientation with a certain angle (θ) in the image [17].

Based on the previous work [17] that proposes textural features must contain fourteen information about texture characteristics. It was found that five best features to be used for analysis and recognition of satellite images are Energy, Contrast, Correlation, Homogeneity, and Entropy of fourteen available features [18]. It means the Energy, Contrast, Correlation, Homogeneity, and Entropy can be used to represent unique information (holistic/global features) of the object type/class.

In this paper, the five features of GLCM are extracted for representing holistic/global information of the cracked or non-cracked image as well as a mild, moderate or severe crack image instead of quantitative metrics (length, width, and area) [2]. Two steps below are the process of extracting cracked or uncracked image features using the GLCM method:

- Creating the GLCM matrix and determining the five GLCM features, namely: firstly, four GLCM matrixes must be generated based on the four GLCM directions(θ) 0°, 45°,90° and 135° [17]. In this case, the distance (d) is set as 1 (one). The GLCM matrix is then normalized so that the sum of all elements is equal to one or each element of the GLCM matrix is divided by the sum of its all elements.
- The GLCM features (Energy, Contrast, Correlation, Homogeneity and Entropy) are calculated from the normalized GLCM normalization by using Eq (1) to Eq. (5) [17]:

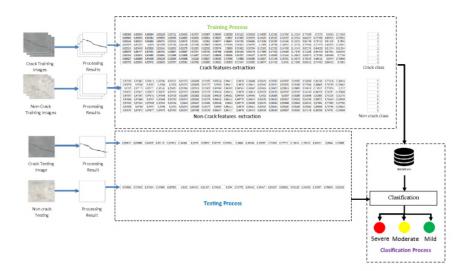


Fig. 1. The block diagram of the crack classification.

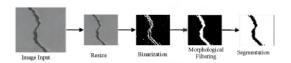


Fig. 2. The sequences of pre-processing for the crack classification.

· Energy or Angular Second Moment (ASM) is interpreted as the quantity of textural uniformity in an image.

$$ASM = \sum_{i} \sum_{j} p(i, j)^{2} \tag{1}$$

· Contrast is represented as the variation moment of a matrix and is the rate of the local irregularities of

$$Contrast = \sum_{n=0}^{Ng-1} n^2 \left\{ \sum_{i=1}^{Ng} \sum_{j=1, |i-j|=n}^{Ng} p(i,j) \right\}$$

· Inverse Difference Moment (IDM) also called degree of homogeneity. If most of the occurrences in GLCM are concentrated close to the main diagonal, then its degree of homogeneity is largevalue.

$$IDM = \sum_{i} \sum_{j} \frac{1}{1 + (i - j)^2} p(i, j)$$
 (3)

· Entropy is represented as the degree of the disorder of an image. The largest value of an entropy will

be obtained if all elements of a matrix are equal.
$$Entropy = -\sum_{i}\sum_{j}p(i,j)\log(p(i,j)) \tag{4}$$

· Correlation is defined as the gray-tone lineardependencies value of an image.

$$ASM = \sum_{i} \sum_{j} p(i, j)^{2} \tag{5}$$

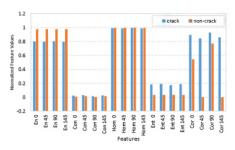
Finally, based on the extraction technique explanation, each image is represented by 20 features obtained from five features of each GLCM orientation matrix. The distribution of GLCM features extracted from two datasets: METU [19]and 177 LE (Crack Dataset from Lombok Earthquake) are presented in Fig. 3(a) and (b). Fig. 3(a) and (b) show that the features of each class appear to be separated from each other, which means that it can be applied to classify cracks and non-cracks image, and level of cracks.

Furthermore, the features are also added with 2 other statistical features, namely the average and standard deviation. Thus each image is totally stated by 22 features.

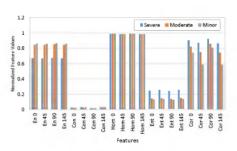
C. Support Vector Machine (SVM) Classifier

The concept of SVM can be explained as an effort to find the best hyperplane that separates two classes in input space. The basic principle of SVM is developed to separate two types/classes of targets, such as fresh and not fresh meat, rain or no rain, etc. Furthermore, SVM can be extended to work on more than two classes problems. In the SVM, the classifications, which can be separated linearly, can be employed by Eq. (6).

$$f(x) = w^T x + b (6)$$



(a) METU dataset



(b) CDLE dataset

Fig. 3. GLCM features distribution.

where $w = \sum_i a_{10}$ and b is a bias, while x is the value of the test image. The best hyperplane parator between the two classes can be found by measuring the hyperplane margin and looking for the maximum point. Margin is the distance between the hyperplane and the closest pattern of each class. This closest pattern is called support vector. The line in Fig. 4 shows the best hyperplane, which is located right in the middle of the two classes while the circle and box located on the dashed line cd and ef are support vectors.

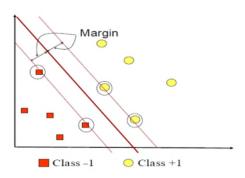


Fig. 4. Example of pre-processing for the crack classification.

The SVM linear classification hyperplane, as given in Fig.

4 and denoted by Eq. 7.

$$w^T x_i + b = 0 (7)$$

By using the SVM classifier model, the input features are classified into class -1, when it meets inequality (Eq. 8).

$$w^T x_i + b < 1 \tag{8}$$

While input features that go into +1 class, is the input features meeting inequality (Eq. 9).

$$w^T x_i + b \ge 1 \tag{9}$$

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

Several experiments were conducted to find out the best model of GLCM features and SVM classifier using two datasets: METU [19] and CDLE. The METU dataset has 40.000 images, which half of them are crack images, and their remaining are non-crack images. The images samples from METU dataset are shown in Fig. 5. The testing on METU dataset aims to evaluate the created SVM model for crack classification, while the testing on CDLE dataset aims to validate whether the crack classification scheme can works correctly on non-standard building crack image due to the earthquake.



Fig. 5. Images samples from METU dataset.

Accuracy, Recall, and Precision metrics are implemented for performance evaluation of the proposed methods.

A. Testing on small size dataset

Firstly, the proposed model was evaluated by using k-fold cross-validation on 1200 and 4800 images set, which were extracted randomly from METU dataset. The two sets of images had an equal amount of data for the type of crack and non-crack, and the ratio between training and testing is 70% versus 30%. The goals of this evaluation were to find the best features model for crack detection and to confirms whether the proposed recognition scheme for crack classification could work properly. The experimental results (see Fig. 6) shows that the GLCM textural features for crack classification must have five information (Energy, Contrast, Correlation, Homogeneity, and Entropy) from four directions (θ : 0° , 45° , 90° and 135°) and distance (d)=1, which is indicated by 98.06% of accuracy for 1200 tested images. This achievement is reproved the previous research [18] that the best GLCM textural features must contain five mentioned information. Additionally, the results also confirm that the proposed method can work properly for crack classification.

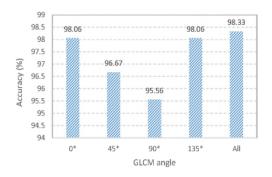


Fig. 6. Accuracy versus GLCM features combinations.

B. Testing on large size dataset

At the second experiment, The proposed scheme was evaluated in a large size dataset (METU has 40.000 images). Additionally, the SVM classifier was compared to the statistical classifiers (Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) [20]). This test intended to determine the robustness of the proposed scheme when applied to a large amount of data and different classifiers. The experimental results (see Fig. 7) show that the proposed method also works properly for large size dataset, which in-line to the achievement of first experiments. In detail, the proposed scheme gives more than 94% of Accuracy, Precision, and Recall, and also shows similar performance to LDA and QDA on METU dataset. It means the proposed scheme for crack classification is powerful enough against a large variety of crack and non-crack images.

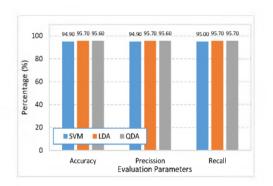


Fig. 7. The performance of the proposed scheme on large size dataset (40000 images) compared to LDA and QDA.

C. Testing on CDLE dataset

CDLE dataset is a crack images collection recorded from damage building due to Lombok earthquake in 2018. The images were captured by using a cellphone camera without capturing standardization in terms of distance the to object. lighting, and resolution. The CDLE had 334 images which were annotated into three types of cracks: mild, moderate, and severe cracks by an expert (Paturrahman, S.T., M.T) from Civil Engineering Dept., Mataram University. Fig. 8 presents the images samples from CDLE dataset.

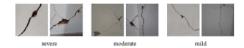


Fig. 8. Images sample from CDLE dataset

Two experiments were carried out using CDLE dataset: testing on two (mild and severe) and three (mild, moderate, and severe) classes. In two classes testing, 56 standard images of 334 images were manually selected based on image sharpness, and the annotation results from an expert. The ratio between training and testing images we 18 et at 90%:10%, which was commonly used for small size data. The experimental results show that the combination of GLCM and SVM give quite good performance indicated averagely about 94.44% of accuracy, 100% of recall, and 88.89% of precision, respectively (see Fig. 9). Additionally, the proposed scheme also achieves much better performances than LDA and QDA classifiers. It means that the combination of GLCM based features with SVM classifier confirms the study of the previous achievements which the GLCM based features are capable of extracting the unique information of the image of building cracks due to the earthquake.

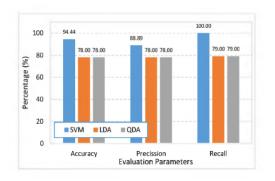


Fig. 9. The performance of the proposed scheme on Lombok dataset for two classes.

The testing for three classes uses 86 images of CDLE dataset with the ratio between data of training and testing: 90%:10%. The last experiment aimed to know the optimality of the proposed algorithm for crack classification. The test results show that the combination of GLCM and SVM give the average accuracy, precision, and recall, about 81.48%, 81.48%, and 88.10%, respectively (see Fig. 10). It can be stated that the combination of GLCM based featurer with SVM classifier is capable of working accurately in the case classification of

the crack image of building due to the Lombok earthquake for three classes (mild, moderate, and severe). It can be

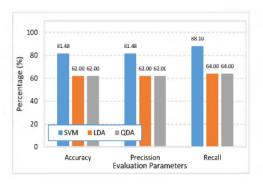


Fig. 10. The performance of the proposed scheme on Lombok dataset for three classes.

achieved, because GLCM features, which is based on second order statistics, indicates the overall average for the degree of correlation among pixel pairs in various perspectives (in terms of homogeneity, uniformity, etc.). Therefore, the GLCM also has been widely implemented for features extraction in any cases and provide high enough accuracy.

V. CONCLUSION AND FUTURE WORK

The scheme of building cracks classification which is based on the GLCM features and SVM classifier works properly indicated by a reasonably high accuracy about more than 94.49 % for the amount of data of 40,000. While on the Lombok earthquake dataset, the proposed scheme provides an average of accuracy, precision, and recall of 94.44 %, 88.89 %, and 100 % respectively for the crack in two-class categories and for the crack in three classes is 81.48 %, 81.48 %, and 88.10 %, respectively. Additionally, the best models of GLCM-based features for building cracks classification are five features, namely: Energy, Contrast, Correlation, Homogeneity and Entropy with LCM angles of 0° , 45° , 90° and 135° .

As the future works, the proposed scheme must be tested by using a large size crack dataset from earthquake to know its robustness against large variability of cracks due to earthquake. In oreder to improve the performance on three classes of the cracks classifusion, an alternative strategy must be developed by using artificial neural networks, convolutional neural networks, or deep-learning.

ACKNOWLEDGEMENT

We send great to thank the Ministry of Research, Technology and Higher Education of the Republic of Indonesia for the research fund under the scheme basic research and contract no:182/SP2H/LT/DRPM/2019. We also thank the expert system laboratory staffs for useful discussions.

REFERENCES

- [1] R. I. Bmkg, "Meteorological, climatological, and geophysical agency of republic of indonesia," 2018. [Online]. Available: https://www.bmkg.go.id/?lang=EN
- [2] S. N. Sheerin, S. Kavitha, and G. Raghuraman, "Review and analysis of crack detection and classification techniques based on crack types," *International Journal of Applied Engineering Research*, vol. 13, no. 8, pp. 6056–6062, 2018.
- [3] A. Mohan and S. Poobal, "Crack detection using image processing: A critical review and analysis," *Alexandria Engineering Journal*, vol. 57, no. 2, pp. 787–798, 2018.
 [4] S. Kabir, "Imaging-based detection of aar induced map-crack damage in
- [4] S. Kabir, "Imaging-based detection of aar induced map-crack damage in concrete structure," NDT & E International, vol. 43, no. 6, pp. 461–469, 2010.
- [5] S. Kabir, P. Rivard, G. Ballivy, and D.-C. He, "Textural analysis for crack-detection using infrared thermography, visual colour, and greyscale concrete imagery," in *Joint International Conference on Computing and Decision Making in Civil and Building Engineering*, 06 2006, pp. 14–16.
 [6] Y. Hu, C.-x. Zhao, and H.-n. Wang, "Automatic pavement crack de-
- [6] Y. Hu, C.-x. Zhao, and H.-n. Wang, "Automatic pavement crack detection using texture and shape descriptors," *IETE Technical Review*, vol. 27, 09 2010.
- [7] V. C. Bharathi and M. K. Geetha, "Segregated handwritten character recognition using glcm features," *International Journal of Computer Applications*, vol. 84, no. 2, 2013.
- [8] J. L. Nunes, M. Piquerez, L. Pujadas, E. Armstrong, A. Fernández, and F. Lecumberry, "Beef quality parameters estimation using ultrasound and color images," *BMC bioinformatics*, vol. 16, no. 4, p. S6, 2015.
- [9] P. Prasanna, K. J. Dana, N. Gucunski, B. B. Basily, H. M. La, R. S. Lim, and H. Parvardeh, "Automated crack detection on concrete bridges," *IEEE Transactions on Automation Science and Engineering*, vol. 13, no. 2, pp. 591–599, April 2016.
- no. 2, pp. 591–599, April 2016.
 [10] Y. Shi, L. Cui, Z. Qi, F. Meng, and Z. Chen, "Automatic road crack detection using random structured forests," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 12, pp. 3434–3445, Dec 2016.
- [11] A. R. Rizvi, P. R. Khan, and S. Ahmad, "Crack detection in railway track using image processing," *International Journal of Advance Research*, *Ideas and Innovations in Technology*, vol. 3, no. 4, pp. 489–496, 2017.
- [12] L. Zhang, F. Yang, Y. D. Zhang, and Y. J. Zhu, "Road crack detection using deep convolutional neural network," in 2016 IEEE international conference on image processing (ICIP). IEEE, 2016, pp. 3708–3712.
- conference on image processing (ICIP). IEEE, 2016, pp. 3708–3712.
 [13] I. Manimozhi and J. Subbiah, "Defect detection in pattern texture analysis using improved support vector machine," Cluster Computing, pp. 1–8, 03 2018.
- pp. 1-8, 03 2018.

 [14] K. Chen, X. Sun, C. Qin, and X. Tang, "Color grading of beef fat by using computer vision and support vector machine," Computers and Electronics in Agriculture, vol. 70, no. 1, pp. 27–32, 2010.
- [15] Z. Hou and J. M. Parker, "Texture defect detection using support vector machines with adaptive gabor wavelet features," in 2005 Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05) -Volume 1, vol. 1, Jan 2005, pp. 275–280.
- [16] R. C. Gonzalez and R. E. Woods, "Digital image processing, volume third edition," 2007.
- [17] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, Nov 1973.
- [18] Z. Lan and Y. Liu, "Study on multi-scale window determination for glcm texture description in high-resolution remote sensing image geo-analysis supported by gis and domain knowledge," ISPRS International Journal of Geo-Information, vol. 7, no. 5, p. 175, 2018.
- [19] Çalar Frat Özgenel, "Concrete Crack Images for Classification," 2018. [Online]. Available: http://dx.doi.org/10.17632/5y9wdsg2zt.1
- [20] A. Tharwat, "Linear vs. quadratic discriminant analysis classifier: a tutorial," *International Journal of Applied Pattern Recognition*, vol. 3, no. 2, pp. 145–180, 2016.

Building Crack Due To Lombok Earthquake Classification Based on GLCM Features and SVM Classifier

ORIGINALITY REPORT

13% SIMILARITY INDEX

11%
INTERNET SOURCES

11%
PUBLICATIONS

6% STUDENT PAPERS

PRIMARY SOURCES

I Wayan Agus Arimbawa, I Gede Pasek Suta Wijaya, Ilham Bintang. "Comparison of simple and stratified random sampling on porn videos recognition using CNN", 2019 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM), 2019

3%

Publication

Publication

Submitted to Wright State University
Student Paper

2%

Eka Mala Sari Rochman, Ika Oktavia Suzanti, Imamah, Muhammad Ali Syakur, Devie Rosa Anamisa, Ach. Khozaimi, Aeri Rachmad. "Classification of Thesis Topics Based on Informatics Science Using SVM", IOP Conference Series: Materials Science and Engineering, 2021

1 %

4

Submitted to Universitas Indonesia
Student Paper

1 %

Exclude quotes Off Exclude matches Off

Exclude bibliography On