Effect of Feature Selection on The Accuracy of Music Genre Classification using SVM Classifier

Conference Paper · September 2020

DOI: 10.1109/iSemantic50169.2020.9234222

8 authors, including:

De Rosal Ignatius Moses Setiadi
Universitas Dian Nuswantoro Semarang

Dewangga Rahardwika
Universitas Dian Nuswantoro Semarang

Eko Hari Rachmawanto
Universitas Dian Nuswantoro Semarang

Atika Sari
Dian Nuswantoro University

Effect of Feature Selection on The Accuracy of Music Genre Classification using SVM Classifier

De Rosal Ignatius Moses Setiadi
Department of Informatics Engineering
Dian Nuswantoro University
Semarang, Indonesia
moses@dsn.dinus.ac.id

Dewangga Satriya Rahardwika
Department of Informatics Engineering
Dian Nuswantoro University
Semarang, Indonesia
rahardwika@gmail.com

Eko Hari Rachmawanto
Department of Informatics Engineering
Dian Nuswantoro University
Semarang, Indonesia
eko.hari@dsn.dinus.ac.id

Christy Atika Sari
Department of Informatics Engineering
Dian Nuswantoro University
Semarang, Indonesia
atika.sari@dsn.dinus.ac.id

Ajib Susanto
Department of Informatics Engineering
Dian Nuswantoro University
Semarang, Indonesia
ajib.susanto@dsn.dinus.ac.id

Ibnu Utomo Wahyu Mulyono
Department of Informatics Engineering
Dian Nuswantoro University
Semarang, Indonesia
ibnu.utomo.wm@dsn.dinus.ac.id

Erna Zuni Astuti
Department of Informatics Engineering
Dian Nuswantoro University
Semarang, Indonesia
erna.zuni.astuti@dsn.dinus.ac.id

Amiq Fahmi
Department of Information System
Dian Nuswantoro University
Semarang, Indonesia
amiq.fahmi@dsn.dinus.ac.id

***Abstract*—This research aims to analyze the effect of feature selection on the accuracy of music genre classification using support vector machine with radial basis function kernel as a classifier. In this research, the music dataset from Spotify is used, which is one of the best-selling music streaming platforms today. The selected feature is metadata because it is considered to have simpler processing than audio feature extraction. The music contained in the Spotify dataset also has complete metadata so that the metadata feature can be used properly. At the feature selection stage, some features are combined in different combination groups (FC1, FC2, FC3, FC4). The classification results prove each feature combination has an accuracy result that has a significant difference, where the best accuracy is 80% and the lowest is 67%. Where the combination of FC1 and FC2 features produces the same accuracy of 80%, but because FC2 has a smaller number of features, so the FC2 combination is recommended because with fewer features, so logically the computing time is shorter.**

Keywords—Feature Selection, Genre, Music, SVM, Spotify

I. Introduction

Music is one of the most important and necessary entertainments in human life. Music becomes a thing that can be used to pour out emotions and human feelings. Today the increasingly massive music industry is developing and being popularized through the internet through applications with streaming platforms, some of which are like Spotify, AppleMusic, SoundCloud, TuneIn, and others. The number of music in the world more and more this is certainly directly proportional to the benefits gained from this industry. This is evidenced by the users one of Spotify's streaming platforms that continue to rise from 68 million in 2015 to 271 million in 2019 [1]. In the music recommendation streaming system application becomes one of the technologies that help users to get information about music that is likely to be liked. The classification of music genres is one of the basic things that are crucial in the music recommendation system [2]. The problem that usually arises is the classification of music genres on high amounts of data and data variations. Classification of music genres has indeed been a challenge in itself because there are no guidelines or standards of what music is like if it is grouped in the same genre, especially in new genres that result from a combination of two or more genres[3]–[5].

Various classification methods such as Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Naive Bayes (NB) have been widely used in various studies [2], [6]–[13]. In research on the classification of music genres, SVM classifiers are more often chosen as in research [2], [6], [9], [10], [12], [14], [15] because SVM is considered relatively more stable and flexible to use with a variety of existing kernels or can also use a custom kernel. SVM has been tested on various datasets such as G. Tzanetakis (GTZAN)[2], [6], [7], [9], [10], Latin Music Dataset (LMD) [9], [14], [16] and International Society of Music Information Retrieval (ISMIR)[9], [14], the results are proven to have stable and satisfying accuracy. As for the KNN method used in the GTZAN dataset in research [10] and NB used for the Garageband dataset in research [3] produces enough accuracy but is not better than SVM. But both of these classifiers are superior in terms of shorter running time than SVM.

In several other studies, a text-based approach has been applied to classify the music genre, for instance in research [17]–[19]. With this approach, music can be classified based on metadata, customer reviews, lyrics, and interpretation. But a text-based approach especially using metadata is considered less effective because the music circulating on the internet may not have complete metadata for various reasons[12]. Even so, the classification of music genres uses metadata relatively faster because metadata can be read directly without the need to extract audio features, all that is needed is the convention of data type only[20]. On music streaming platforms such as Spotify music genre classification is possible based on metadata, because the metadata in the music dataset is relatively very complete. Spotify Music Dataset that contains 228,159 metadata from a music that is unique to each other and is grouped into 26 different genres and has 18 features of metadata in each music[21]. Of the 18 metadata features found in the Spotify music dataset, there are genre metadata. In this research, genre metadata will not be used in the classification process but is used as a target and determines classification accuracy. Because the selection of the number of features is very relatively large, there is no need for a classification process to use all available features. This is because the number of features also affects the computation and complexity of the classification process[22]. Then in this research will be tested and analyzed the effect of the number of features on the classification accuracy of the SVM classifier.

II. METHOD

In this section the stages of the proposed research are explained so that it is easier to understand can see Fig.1. Based on Fig.1 it can be explained as follows.

1. Music Dataset Acquisition: Spotify Music Dataset downloaded from the page www.crowdai.org contains 26 music genres with a total of 228,159 music with 18 features in the form of acquired metadata.
2. Selecting several genres: From a total of 26 existing music genres, 5 genres are selected, with 6,000 music data taken from each genre.
3. Separating data into datasets and targets: This data separation aims to facilitate the training process. The data is divided into two, namely the dataset which contains the entire data used with 17 metadata features except for genre features. While the target is a genre feature that can be used for the labeling process.

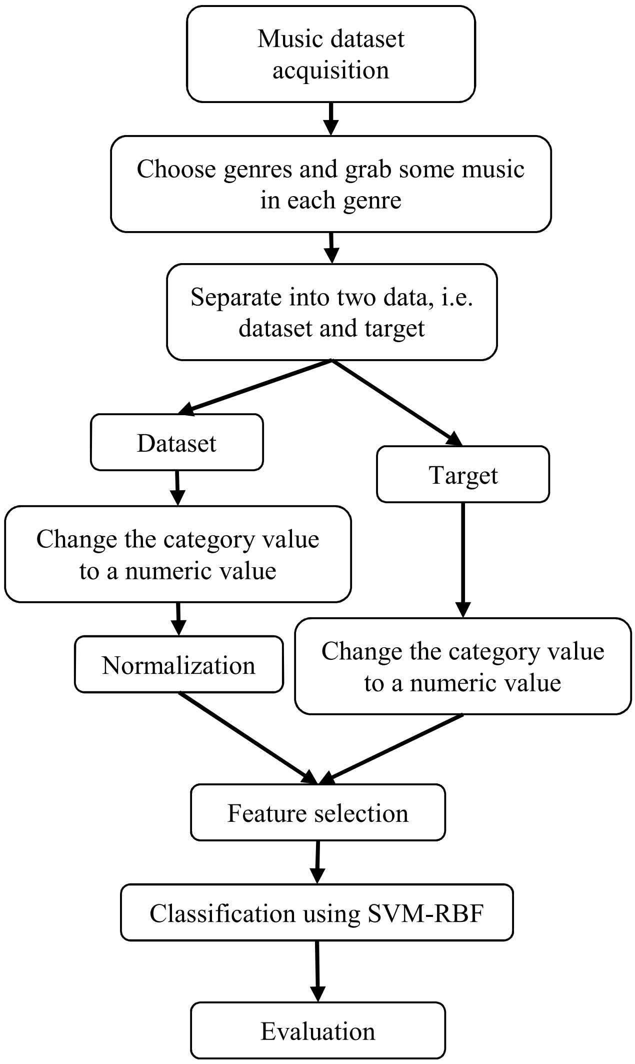

Fig. 1. The stages of the method used in research

4. Changing categorical value: Because the value of the feature is not numeric, but the value in the form of an object or categorical, then the conversion of categorical values into numerical values can be performed mathematical calculations.
5. Normalization: the process of normalizing data is useful for narrowing the range of data. Data is transformed in the range of 0 to 1 with a standard deviation of 1.00 and a mean value of 0.00 because SVM can work more optimally if all feature values are centered at 0.00 and have the same variant value.
6. Features Selection: Because there are a large number of metadata features, a selection of features that are considered most important, features that are considered not important are not used. The reduction in the number of features aims to speed up the computing process. In the next section, the classification results will be tested using some different features, namely 5, 10, 10, and 13, respectively.
7. Classification: at this stage, SVM-RBF is proposed as a classifier. Based on the dataset used, it will be divided into two parts, namely training data and testing data with a ratio of 80% and 20%.
8. Evaluation: this stage is used to measure the accuracy value by comparing the actual value and the predicted value.

III. IMPLEMENTATION AND TESTING

In this research, the Spotify music dataset was chosen as the music dataset used. This dataset is a music dataset that has 228,159 music with 26 genres and 18 features [21]. To process it, the Python programming language is used, besides that, it uses the Python Data Analysis Library[23] to process data structures and perform data analysis and Scikit Learn to support the classification process. Scikit Learn is a package that contains important modules that can support machine learning projects[24]. In detail the following stages of research following the methods described in the previous section.

A. Choose genres and grab some music

A large number of music and features in the Spotify dataset makes this dataset has many advantages, where researchers can use all or part of the dataset to conduct experiments. Because the number of music in each genre is not the same, then in this research, not all music is used but the same amount is taken in each dataset. Because of the huge number of data and the limited hardware capabilities, this research selected five music genres, where each genre was taken from 6000 music. The five genres are Pop, Electric, Rap, Opera, and Folk, so the total music used is thirty thousand.

B. Separate Dataset

The selected music data is then divided into datasets and targets. Dataset is data with metadata used in the process of implementing the method, while the target is the genre that is owned by music. Furthermore, music data is divided into training data and testing data with each experiment having a ratio of 80% training data and 20% testing.

C. Change the category value to a numeric value

At this stage, the process is carried out using PANDAS. This is done so that the mathematical calculation process can be done. Because the data types of metadata are quite diverse, some of which are object data types.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 228159 entries, 0 to 228158
Data columns (total 18 columns):
genre               228159 non-null object
artist_name         228159 non-null object
track_name          228159 non-null object
track_id            228159 non-null object
popularity          228159 non-null int64
acousticness        228159 non-null float64
danceability        228159 non-null float64
duration_ms         228159 non-null int64
energy              228159 non-null float64
instrumentalness    228159 non-null float64
key                 228159 non-null object
liveness            228159 non-null float64
loudness            228159 non-null float64
mode                228159 non-null object
speechiness         228159 non-null float64
tempo               228159 non-null float64
time_signature      228159 non-null object
valence             228159 non-null float64
dtypes: float64(9), int64(2), object(7)
memory usage: 31.3+ MB
```

(a)

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 30000 entries, 12708 to 76828
Data columns (total 17 columns):
artist_name         30000 non-null float64
track_name          30000 non-null float64
track_id            30000 non-null float64
acousticness        30000 non-null float64
danceability        30000 non-null float64
duration_ms         30000 non-null float64
energy              30000 non-null float64
instrumentalness    30000 non-null float64
key                 30000 non-null float64
liveness            30000 non-null float64
loudness            30000 non-null float64
mode                30000 non-null float64
speechiness         30000 non-null float64
tempo               30000 non-null float64
time_signature      30000 non-null float64
valence             30000 non-null float64
popularity          30000 non-null float64
dtypes: float64(17)
memory usage: 4.1 MB
```

(b)

Fig. 2. Figure.2 Converting features value results{(a)Before conversion; (b)After conversion}

The mathematical calculation process requires numeric data types (float or integer), because the features possessed do not all have numeric data types, it is necessary to convert the categorical value to a numeric data type. The conversion process can be done with the label encoding process; the result of the encoding label will make the feature value change to a numeric data type so that it can be calculated with the SVM algorithm. To see in more detail this stage, see Fig.2.

D. Normalization

The data normalization process is carried out to equalize the range of data values that differ from one another. In this study, the authors change these values to the range of 0.00-1.00. This is done because the SVM can calculate correctly and optimally in that range. As an example can see Fig.3.

	acousticness	instrumentalness	energy	loudness
count	30000.000000	30000.000000	30000.000000	30000.000000
mean	0.383267	0.130487	0.537513	-9.834656
std	0.377461	0.273664	0.268219	6.288355
min	0.000004	0.000000	0.001850	-45.539000
25%	0.035300	0.000000	0.324000	-12.552250
50%	0.220000	0.000118	0.574000	-7.758500
75%	0.806000	0.043700	0.753000	-5.475000
max	0.996000	0.994000	0.999000	1.585000

(a)

	acousticness	instrumentalness	energy	loudness
count	30000.000000	30000.000000	30000.000000	30000.000000
mean	0.384803	0.131275	0.537194	0.757668
std	0.378978	0.275316	0.268986	0.133443
min	0.000000	0.000000	0.000000	0.000000
25%	0.035438	0.000000	0.323071	0.699999
50%	0.220881	0.000119	0.573785	0.801725
75%	0.809236	0.043964	0.753297	0.850182
max	1.000000	1.000000	1.000000	1.000000

(b)

Fig. 3. Figure.3 Sample of normalization results {(a) Before Normalization; (b) After Normalization}

E. Features Selection

Spotify Music Dataset provides 17 features in addition to the genre, in the form of metadata that can be used for various processes. However, the selection of features can be crucial given the importance of these features can affect the classification results. So, feature selection becomes a very important thing to do. Feature selection can be done by removing and or using certain features to optimize running time and even improve the accuracy of the methods used. The feature selection process uses the chi-square calculation by sorting the relationship of a feature with the target class in the dataset used, see Table 1 for the results of the calculations.

TABLE I. FEATURES SCORE

No	Features	Score
1	Acousticness	7333.825205
2	Instrumentalness	3664.597613
3	Popularity	2780.085255
4	Energy	2314.629554
5	Danceability	1032.640592
6	Mode	1048.192344
7	Speechiness	1032.152780
8	Valence	922.493163
9	Loudness	451.282083
10	Duration_ms	137.867351
11	Tempo	117.405865
12	Liveness	88.828267
13	Artist_name	86.447100
14	Time_signature	36.059803
15	Track_name	11.493590
16	Key	10.568912
17	Track_id	5.321589

TABLE II. FEATURES COMBINATION USED

Name	Features
FC1	Acousticness, instrumentalness, popularity, energy, danceability, speechiness, valence, loudness, tempo, liveness, artist_name, time_signature, key
FC2	Acousticness, instrumentalness, popularity, energy, danceability, speechiness, valence, loudness, tempo, artist_name
FC3	Acousticness, instrumentalness, popularity, energy, danceability
FC4	Acousticness, instrumentalness, popularity, energy, danceability, speechiness, valence, loudness, duration_ms, mode

The higher score value in Table 1 indicates that the importance of a feature in the target classIn this research four feature combination combinations are also used which are presented in Table 2. From Table 2, there can be seen four combinations where the FC1 feature combination contains 13 features out of 17 features with the highest score and eliminates four other features namely duration_ms, mode, track_name, and track_id. FC2 is taken from 10 features with the highest score by removing duration_ms and mode features, then adding artis_name and tempo features instead, while in FC3 and FC4 5 and 10 features are taken with the highest value. The consideration of adding and removing features in FC2 is based on logic, where the artist's name and tempo greatly affect the classification, especially the name of the artist is one of the keywords to search for new music in the real world.

F. *Classification using SVM-RBF and Evaluation*

In this research, the SVM classifier with the RBF kernel is used. In the RBF kernel, there is a gamma variable (γ) that is used to explain how far the influence of each data in determining the decision boundary. Low gamma value means that individual data is far from the decision boundary otherwise, high gamma value means the data is near the decision boundary which can affect the decision boundary. Besides, there are also variable costs (C). This variable is a hyperparameter that can determine how many penalties are given to the data so that it can determine the margin used. A high C value is often referred to as a hard margin which means the margin does not provide a penalty for misclassification data so that it will narrow the margin value. While a low C value can be called a soft margin that can expand margins on decision boundaries. Values from Gamma and C cannot be given with only one fixed value. So we need to make a range value for the variables Gamma and C. In this research Gamma_range between 0.0001 to 1 and C_range of 0.01 to 100.

After selecting the value, the SVM-RBF classification process is carried out with Grid Search. Grid Search can help to find the most optimal results from the use of hyperparameter that was previously used. Grid Search provides the highest accuracy output by combining the hyperparameter. However, it must first be done to determine the value of cross-validation first because the calculations carried out SVM based on cross-validation. The author uses k-fold cross-validation with a comparison of test data of 20% and training data of 80% with a free random state. The random state value is equalized for each feature combination. After that, then the accuracy calculation will be the output of Grid Search which becomes a benchmark for the performance evaluation method. The results of the classification of five genres by the proposed method are presented in Table 3.

TABLE III. CLASSIFICATION RESULTS IN FIVE GENRES

Features combination	C	γ	Accuracy
FC1	100	0.1	80%
FC2	100	1.0	80%
FC3	100	1.0	76%
FC4	10	1.0	67%

C value and Gamma value are hyperparameters that were previously only being analyzed with minimum and maximum values that can be used. The value that appears on the C value and Gamma value in table 3 is a combination of C and Gamma values that can produce the most optimum accuracy in the Grid Search. The highest accuracy value is shown in experiments using a combination of FC1 and FC2 features with the same accuracy that is 80%. The lowest results and lags far behind other experiments are the combination of FC4 features by producing an accuracy value of 67%.

The accuracy results presented in Table 3 show that feature selection has a significant effect on accuracy. Feature selection cannot be done arbitrarily; feature selection must be done correctly following scientific theory. However, the above experiments show that the combination of FC1 and FC2 features have similar results, but seen from the number of features used there are differences. Where FC1 uses thirteen features and FC2 uses ten. The combination of FC1 and FC2 features has the same features, it's just that in FC2 features liveness, time_signature, key. Because the number of features is smaller than it can reduce the computational time needed.

IV. CONCLUSIONS

Classification is one of the methods found in data mining. In the classification, there are several main stages namely, preprocessing, feature extraction and selection, training, and testing. The process of extracting and selecting features is one of the keys to success in improving classification accuracy. Previous research has found several findings which state that the use of all feature extractions does not necessarily increase the level of classification accuracy, but instead burdensome the computational process due to more features that must be calculated. In this research, a trial was conducted with various feature selection options in the classification of music genres. A large number of features available are indeed very pronounced if used entirely. From some feature selection, there is a feature selection option with chi-square, as well as a few manual modifications based on logic. In FC1 13 features are used out of 17 features taken based on the highest chi-square value, FC2 has selected 10 features out of 17 features taken based on the highest chi-square value, but changing duration_ms and mode features with artis_name and tempo features, FC3 and FC4 use 5 and the 10 highest features based on the chi-square sequence. It turned out that the highest accuracy was obtained by FC1 and FC2 with an accuracy of 80%. however, the authors recommend using FC2 which has fewer features, because it can reduce the computational time required. With these results, it can be concluded that in classification research it is necessary to choose features correctly to produce optimal accuracy and faster computing time.

REFERENCES

[1] A. Watson, "Spotify's monthly active users 2015-2019," *Statista*, 2019. [Online]. Available: https://www.statista.com/statistics/367739/spotify-global-mau/. [Accessed: 02-Feb-2020].

[2] A. C. M. da Silva, M. A. N. Coelho, and R. F. Neto, "A Music Classification model based on metric learning applied to MP3 audio

files," *Expert Syst. Appl.*, vol. 144, Apr. 2020.

[3] M. M. Panchwagh and V. D. Katkar, "Music genre classification using data mining algorithm," in *Conference on Advances in Signal Processing, CASP 2016*, 2016, pp. 49–53.

[4] B. Zhang and J. Lin, "An efficient content based music retrieval algorithm," in *Proceedings - 3rd International Conference on Intelligent Transportation, Big Data and Smart City, ICITBS 2018*, 2018, vol. 2018-Janua, pp. 617–620.

[5] Adiyansjah, A. A. S. Gunawan, and D. Suhartono, "Music recommender system based on genre using convolutional recurrent neural networks," in *Procedia Computer Science*, 2019, vol. 157, pp. 99–109.

[6] P. Fulzele, R. Singh, N. Kaushik, and K. Pandey, "A Hybrid Model for Music Genre Classification Using LSTM and SVM," in *2018 11th International Conference on Contemporary Computing, IC3 2018*, 2018.

[7] S. Sharma, P. Fulzele, and I. Sreedevi, "Novel hybrid model for music genre classification based on support vector machine," in *ISCAIE 2018 - 2018 IEEE Symposium on Computer Applications and Industrial Electronics*, 2018, pp. 395–400.

[8] G. Kour, M. Tech, S. Rbiebt, N. Kharar, and A. Mehan, "Music Genre Classification using MFCC, SVM, and BPNN," 2015.

[9] L. Nanni, Y. M. G. Costa, A. Lumini, M. Y. Kim, and S. R. Baek, "Combining visual and acoustic features for music genre classification," *Expert Syst. Appl.*, vol. 45, pp. 108–117, Mar. 2016.

[10] D. P. Kumar, B. J. Sowmya, Chetan, and K. G. Srinivasa, "A comparative study of classifiers for music genre classification based on feature extractors," in *2016 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics, DISCOVER 2016 - Proceedings*, 2016, pp. 190–194.

[11] N. D. Al-Shakarchy, H. F. Al-Shahad, and D. A. Al-Nasrawi, "Cryptographic system based on Unicode," in *Journal of Physics: Conference Series*, 2018, vol. 1032, no. 1.

[12] A. J. H. Goulart, R. C. Guido, and C. D. MacIel, "Exploring different approaches for music genre classification," *Egypt. Informatics J.*, vol. 13, no. 2, pp. 59–63, Jul. 2012.

[13] O. R. Indriani, E. J. Kusuma, C. A. Sari, E. H. Rachmawanto, and D. R. I. M. Setiadi, "Tomatoes classification using K-NN based on GLCM and HSV color space," in *Proceedings - 2017 International Conference on Innovative and Creative Information Technology: Computational Intelligence and IoT, ICITech 2017*, 2018, vol. 2018-Janua, pp. 1–6.

[14] Y. M. G. Costa, L. S. Oliveira, A. L. Koerich, F. Gouyon, and J. G. Martins, "Music genre classification using LBP textural features," *Signal Processing*, vol. 92, no. 11, pp. 2723–2737, Nov. 2012.

[15] A. R. Rajanna, K. Aryafar, A. Shokoufandeh, and R. Ptucha, "Deep neural networks: A case study for music genre classification," in *Proceedings - 2015 IEEE 14th International Conference on Machine Learning and Applications, ICMLA 2015*, 2016, pp. 655–660.

[16] A. L. Przybysz, R. Corassa, C. L. D. Santos, and C. N. Silla, "Latin Music Mood Classification Using Cifras," in *Proceedings - 2015 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2015*, 2016, pp. 1682–1686.

[17] S. Oramas, L. Espinosa-Anke, A. Lawlor, X. Serra, and H. Saggion, "Exploring Customer Reviews for Music Genre Classification And Evolutionary Studies," in *Proceedings of the 17th ISMIR Conference*, 2016, pp. 150–156.

[18] R. Mayer, R. Neumayer, and A. Rauber, "Rhyme and style features for musical genre classification by song lyrics," in *ISMIR Conference*, 2008.

[19] K. Choi, J. H. Lee, and J. S. Downie, "What is this song about anyway?: Automatic classification of subject using user interpretations and lyrics," in *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, 2014, pp. 453–454.

[20] X. Serra, "Audio content processing for automatic music genre classification: descriptors, databases, and classifiers. Enric Guaus i Termens."

[21] B. Brost, R. Mehrotra, and T. Jehan, "The Music Streaming Sessions Dataset," 2018.

[22] C. Irawan, E. N. Ardyastiti, D. R. I. M. Setiadi, E. H. Rachmawanto, and C. A. Sari, "A Survey: Effect of the Number of GLCM Features on Classification Accuracy of Lasem Batik Images using K-Nearest Neighbor," in *International Seminar on Research of Information Technology and Intelligent Systems*, 2018, pp. 33–38.

[23] W. McKinney, *Python for data analysis : data wrangling with pandas, NumPy, and IPython*, 2nd ed. O'Reilly Media, Inc, 2018.

[24] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.