# Data Segmentation and Concatenation for Controlling K-Means Clustering-Based Gamelan Musical Nuance Classification

Heribertus Himawan[1], Arry Maulana Syarif[2], Ika Novita Dewi[3], Abdul Karim[4]
Computer Science in Arts and Culture Research Center, Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang, Indonesia[1, 2, 3]
Cerebrovascular Disease Research Center, Department of Artificial Intelligence Convergence, Hallym University, Republic of Korea[4]

***Abstract*—The musical nuance classification model is proposed using a clustering-based classification approach. Gamelan, a traditional Indonesian music ensemble, is used as the subject of this study. The proposed approach employs initial and final data segmentation to analyze symbolic music data, followed by concatenation of the clustering results from both segments to generate a more complex label. Structural-based segmentation divides the composition into an initial segment, representing theme introduction, and a final segment, serving as a closing or resolution. This aims to capture the distinct characteristics of the initial and final segments of the composition. The approach reduces clustering complexity while maintaining the relevance of local patterns. The clustering process, performed using the K-Means algorithm, demonstrates strong performance and promising results. Furthermore, the classification rules derived from data segmentation and concatenation help mitigate clustering complexity, resulting in an effective classification outcome. The model evaluation was conducted by measuring the similarity within the classes formed from data merging using Euclidean distance score, where values below three indicate high similarity, and values greater than ten indicate strong dissimilarity. Three of the 13 formed classes with more than one data point, Class 5, Class 12, and Class 18, demonstrate high similarity with a value below three. Five other classes, Class 7, Class 10, Class 11, Class 15, and Class 20, exhibit near-high similarity, with values ranging from three to four, while the remaining five classes fall within the range of four to five.**

***Keywords—Musical emotion clustering; classification; clustering-based classification; K-Means algorithm; symbolic music; gamelan music***

## I. Introduction

Musical Emotion Classification (MEC) aims to group compositions into emotional categories such as joy, anger, sadness, or calmness, [1]. MEC research has been expanding due to the increasing importance of mood in music applications, beyond just musical genres [2]. MEC plays a crucial role in music recommendation systems [3], psychotherapy, and music visualization [4]. The exploration of musical emotions in music psychology and music information retrieval research on sacred music remains largely unexplored [5]. This condition also applies to traditional music, such as Gamelan, a traditional musical form from Java, Indonesia.

Melodies in Western music have musical emotion characteristics that can be identified through tempo, dynamics, major or minor scale modes, and other elements. Meanwhile, although gamelan music pieces have different musical emotions, they are played with similar techniques and tempos. Therefore, MEC in gamelan music requires a different approach. This study proposes a novel mathematical approach using a clustering-based classification method to classify musical emotions in gamelan music, a genre characterized by a high degree of similarity among different musical emotions. The K-Means clustering-based musical emotion classification model introduced in this study presents a novel approach to solving the problem of label-free datasets, where no predefined emotional labels exist for each composition. The melodic sequence dataset is treated as categorical nominal data rather than ordinal data. Since K-Means clustering is well-suited for categorical nominal data, it is chosen as the clustering algorithm for this study. Data segmentation and concatenation methods are used to control the K-means algorithm in performing clustering, where the composition data is segmented, and then data concatenation is used to determine the musical emotions class based on the cluster output of the data segment.

The clustering output consists of numerically labeled clusters, where each cluster contains compositions that share similar musical emotions, without explicitly describing the specific type of emotional expression. Due to the lack of clear and validated reference sources on gamelan compositions categorized by musical emotions, the term musical nuance classification is more appropriate in this context. In other words, musical nuance classification aims to group compositions based on mathematical similarity in melodic patterns. The choice of the term musical nuance also serves as a form of respect toward the gamelan community, acknowledging that musical emotions in gamelan music remain undefined and debated. The main novelty of our approach focuses on interpreting musical emotions through the generation of more complex musical emotion labels based on segmentation and concatenation of data taken from the beginning of the melody to represent the introduction of the theme, and the end of the melody to represent the ending or resolution.

The main novelty of our approach focuses on the generation of more complex musical emotion labels based on segmentation and concatenation of data taken from the beginning of a melody to represent the introduction of a theme, and the end of a melody to represent the ending or resolution. The results of this research can contribute to composition selection for datasets at the level of musical nuance, supporting applications such as automatic music generation systems and music recommendation systems. The details of the proposed model are structured as follows: Section II reviews relevant MEC research. Section III explains the methodology used in this study. Sections IV and V present the experimental results and discussion, respectively. Finally, Section VI provides the conclusions drawn from the proposed MEC model.

## II. Related Work

The incorporation of musical emotions in music generation is achieved using a supervised learning approach, where the system is trained on labeled data categorized by musical emotion classes [6]. The output from the MEC was used as the basis for automatic music generation through emotion-based composition selection performed by study [7]. At the low level, MEC is performed by processing audio signals using features such as spectrum, rhythm, and mel-frequency cepstrum coefficients (MFCCs) [1]. At the high level, MEC analyzes relationships between musical elements within a composition, such as pitch, note duration, and rhythm [2]. Supervised learning is a widely adopted approach in MEC, where emotional class labels are assigned to compositions in a dataset, as demonstrated in the development of an MEC system using Convolutional Neural Networks (CNN) by study [8], and the development of a clustering system for personalized music labels using a tag-based collaborative filtering algorithm by study [9]. However, some musical traditions, particularly traditional music, require unsupervised learning approaches to classify compositions based on emotion. The supervised learning approach using CNN has been applied in the development of MEC systems for Indian traditional music, which is known for its ambiguous nature [10]. Ambiguity is also a characteristic of gamelan music, a traditional musical ensemble from Java, Indonesia, where the definition of musical emotion classes remains a subject of academic discussion.

Unlike Western music, where specific emotions are often embedded within a song, emotions in gamelan music are highly dependent on the tempo and playing style of the musicians. Changes in tempo within a single performance are common [11]. Meanwhile, the study in [12] describes that "*Rasa*, in a Javanese musical context, has many meanings that range from affect, feeling, and inner meaning to perception, understanding, and intuition." He defines musical emotions using the term *rasa*, which can be simply translated as feeling in Javanese. *Rasa* expresses emotions that evoke sensations and attempts to formulate rules related to musical emotions in gamelan music. However, such formulations tend to remain within academic discussions of gamelan music and are not widely recognized by gamelan musicians. The study in [13] stated that the *rasa* of a melody can change depending on how the musician plays the music. This is similar to jazz music, where the interpretation of a song's emotion can vary based on the musician's performance.

The classification of musical emotions in gamelan music remains a subject of debate. However, there are compositions that are traditionally played at specific ceremonial events. For example, "Kebo Giro", an instrumental piece, is performed during wedding ceremonies, while "Suwe Ora Jamu" is a well-known song with cheerful lyrics. Based on these observations, this study assumes that emotional classes exist in gamelan compositions. However, for reasons not fully agreed upon by the gamelan community, these emotional classes are not explicitly defined within gamelan compositions. This is evident from the absence of datasets categorizing gamelan compositions based on musical emotions. Even when such data exists, it consists of subjective interpretations from different individuals and is highly limited. Therefore, rather than using a supervised learning approach, musical emotion classification in Gamelan music is more appropriately conducted using a clustering-based classification approach, where composition data lacks predefined musical emotion labels.

In MEC, a non-masked language model from Natural Language Processing (NLP) methods is used for pre-training large-scale unlabeled music data, followed by fine-tuning on the pre-trained model [14]. The deductive approach is an interesting avenue for exploration in MEC, particularly for traditional music, although it may require additional effort for fine-tuning. On the other hand, an inductive approach is also a logical direction, where experiments are conducted on specific traditional music, and the results are further developed for Western music, a more widely recognized genre. Therefore, the novelty of models and methods in research on music classification, automatic music generation, and related fields conducted on traditional music can serve as a reference for generalization to more common and popular music genres.

Feature selection based on musical notation, including the recognition of its physical characteristics, is essential for the development of MEC systems. Data segmentation is one of the methods used to optimize classification time [15]. Feature selection in gamelan music was conducted by [16] by calculating the odd-even positions of notes within a melodic sequence. Meanwhile, [17] represented gamelan musical rules, such as beat, meter, pitch variations, and note duration, as features for an LSTM-based gamelan music generation system. The concept of musical pattern balance through quantification of notes based on odd-even sequencing, as proposed by [16], is an intriguing approach for further exploration. Potential applications include representing the beginning and ending of melodies as an introduction and resolution, analyzing the balance of variance and note distribution in musical patterns, and other related aspects. Although metadata, such as artist name, album, genre, and other attributes, can be used for musical emotion classification, content-based feature analysis based on musical elements is more adaptive in identifying musical emotions and listener preferences [18].

In symbolic music classification, compared to rule-based models that rely on presumptions based on predefined rules, data-driven models, which make assumptions based on statistical analysis of a sequence of events for grouping perception, are more robust in inferring simple and concise text-based musical elements. Although deep learning and NLP methods have proven to be reliable in automatic music generation and MEC, AI methods are still needed to better control certain musical features that are difficult to recognize using deep learning alone [19]. The use of symbols can be a solution to issues of richness and ambiguity in natural language by providing a simple and unique representation of data [20]. On the other hand, hybrid methods combining deep learning (DL) with AI or machine learning (ML) with rule-based methods can enhance system performance. Musical emotion classification using Gated Recurrent Unit (GRU) has been integrated with structural music analysis through a rule-based method [21]. The rule-based method can be utilized to analyze musical rules that serve as constraints in MEC or act as a decision-maker for classification or clustering outputs generated by ML, NLP, or DL methods.

The K-Means algorithm has been widely used in MEC research to address the challenge of unlabeled musical data. Studies have applied K-Means for various purposes, such as musical emotion classification in opera music, where compositions were grouped into four types of musical emotions using unsupervised datasets [22]. Other applications include the conversion of unstructured musical data into structured music features [23] and music recommendation systems that consider both musical data and user preferences in large-scale music datasets [24]. In this study, K-Means clustering is applied to a small dataset containing 49 Gamelan compositions. Although scalability is not the primary focus, the findings are expected to highlight the potential of the proposed model for generalization to larger datasets and various musical genres, making it an alternative goal worth exploring. In order to clarify our research position, a brief summary of the selected MEC research is presented in Table I.

TABLE I. BRIEF SUMMARY OF SELECTED MEC RESEARCH

| Research | Model | | Melodies | Dataset | Task |
|---|---|---|---|---|---|
| Qiu et al. [14] | Supervised-learning | Transformer - Deep Learning | 1071 | EMOPIA dataset | Learning musical emotions based on sequence-level classification and note-level classification |
| | | | 7191 | VGMIDI dataset | |
| Lian [15] | Supervised-learning | Radial Basis Function Neural Networks | 1608 | AMG1 608 dataset | Classifying musical emotions based on the Thayer and Hevner emotion models |
| Jia [5] | Supervised-learning | CNN-LSTM | 5286 | Chinese audio and lyrics | Classifying musical emotions based on lyrics and note sequences |
| Ferreira et al. [19] | Unsupervised-learning | Transformer - Deep Learning, and Monte Carlo Tree Search | 728 | VGMIDI dataset | Controlling musical emotions using Monte Carlo Tree Search for symbolic music generation |
| | | | 3122 | NinSheetMusic community | |
| Chaudhary et al. [8] | Supervised-learning | CNN - Deep Learning | 1000 | Hindi songs | Classifying musical emotions based on music spectrogram signals |
| Medina et al. [2] | Supervised-learning | Multilayer Perceptron | 1802 | MediaEval dataset | Classifying musical emotions based on dimensions that describe the emotional qualities of music: valence and arousal |
| Ours | Unsupervised-learning | K-Means Clustering | 49 | Gamelan Music Scores | Interpreting musical emotions based on data segmentation and data concatenation taken from the beginning and the end of melodies. |

## III. METHOD

Gamelan music consists of two types of musical scales: laras pelog and laras slendro. Laras pelog comprises seven pitches: 1, 2, 3, 4, 5, 6, and 7, while laras slendro consists of five pitches: 1, 2, 3, 5, and 6. The pitches in these two musical scales have different audio frequencies. In addition to these notes, there is a rest note, which represents a moment of silence. To facilitate computational processing and simplify notation, the rest note is converted into the number 0. Each musical scale contains three musical modes, characterized by the dominance of specific pitches within a composition. Laras pelog consists of pathet barang (pelog barang), pathet lima (pelog lima), and pathet nem (pelog nem), while laras slendro consists of pathet manyura (slendro manyura), pathet nem (slendro nem), and pathet sanga (slendro sanga). Although both laras pelog and laras slendro have a musical mode called pathet nem, their compositional characteristics differ. Fig. 1 illustrates the structure of Gamelan music based on its musical scale and musical mode.

Clustering is performed on musical symbolic data in the form of note sequences. Segmentation is applied to compositions by extracting a portion of the note sequence from the beginning of the composition and another portion from the end. A composition follows a plot or storyline, represented by the movement of note sequences from start to finish. The movement of note sequences can be identical or different at the beginning and the end, or they may start similarly but diverge towards the end, and vice versa. Furthermore, even if two compositions exhibit the same note sequence movement pattern, differences in their movement characteristics may still exist. It should be underlined that the proposed method is limited to melodic data that has the same structure as the data used in this experiment, where each beat contains one note.

The clustering process is conducted separately on the note sequence segments at the beginning, referred to as the initial segment, and at the end of the composition, referred to as the final segment. Differences in the musical nuance classes between the initial and final segments within a single composition are possible. The clustering outputs from both segments are then used as input to determine the musical nuance class of each composition using a data concatenation technique. Fig. 2 illustrates the data segmentation and concatenation model in K-Means clustering-based Gamelan musical nuance classification. In general, the research method consists of three stages: data collection and pre-processing, clustering-based classification, and evaluation.

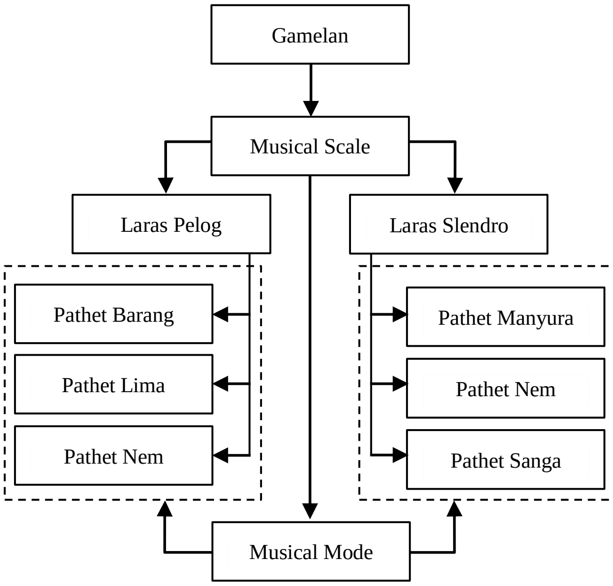


Fig. 1. Illustration of gamelan music structure based on musical scale and musical mode.

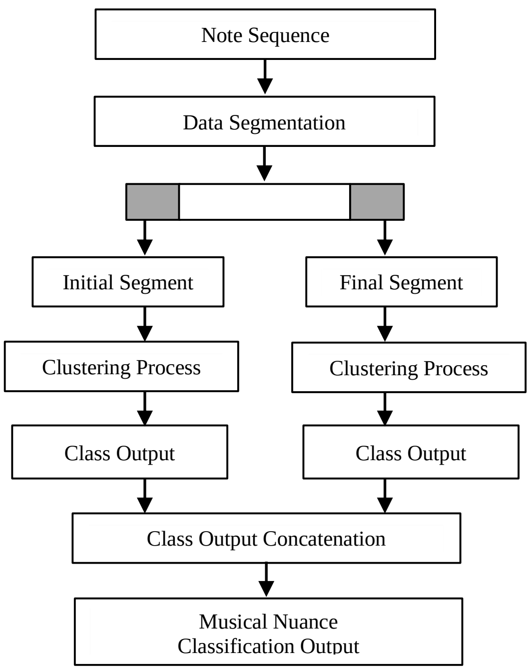


Fig. 2. Illustration of the data segmentation and concatenation model in K-Means clustering-based gamelan musical nuance classification.

### A. Data Collection and Pre-Processing

The dataset consists of notation sequences from Gamelan music, collected from various Internet sources and literature. The collected data includes 49 compositions in the Pelog Barang scale. The notation sequences are derived from the skeletal melody, which functions similarly to chords in Western music. A Gamelan composition consists of metrical sequences, where each metrical unit (bars) contains four beats or notes. These bars are arranged into metrical rows (rows), where each row consists of two bars. Fig. 3 illustrates an example of a *pelog barang* composition titled "Biwadha Praja", which consists of 24 bars.

Biwadha Praja
(Laras Pelog Pathet Barang)

| | |
|---|---|
| 2 2 0 0 | 2 2 0 3 |
| 5 5 6 5 | 3 5 6 7 |
| 0 7 6 5 | 3 5 7 6 |
| 7 5 6 7 | 6 5 3 2 |
| 3 2 7 6 | 5 6 7 2 |
| 3 3 2 7 | 6 5 3 2 |

Fig. 3. A sample of Pelog Barang gamelan composition titled 'Biwadha Praja' used in the experiment.

The notation sequence data from each gamelan composition is transformed into an array format for computational processing. Using the example composition from Fig. 1, the notation sequence is (2, 2, 0, 0, 2, 2, 0, 3, 5, 5, 6, 5, 3, 5, 6, 7, 0, 7, 6, 5, 3, 5, 7, 6, 7, 5, 6, 7, 6, 5, 3, 2, 3, 2, 7, 6, 5, 6, 7, 2, 3, 3, 2, 7, 6, 5, 3, 2). The transformation of notation sequences into array format is applied to all compositions. Given P as the set of notation sequences, we define P = ($p_1$, $p_2$, $p_3$, ..., $p_n$). Since not all compositions have the same number of metrical units (birama), segmentation is performed to standardize vector lengths. For example, the composition "Biwadha Praja" consists of 24 birama, while another *pelog barang* composition, "Asmaradana," consists of only eight bars: (2, 7, 2, 6, 2, 7, 2, 3, 5, 3, 2, 7, 3, 2, 3, 7, 6, 3, 2, 7, 3, 2, 7, 6, 5, 3, 2, 7, 3, 2, 7, 6). To address this variation, data segmentation is applied by dividing the notation sequence into three segments: the initial segment representing the early part of the composition, the body segment representing the middle section, and the final segment representing the ending part of the composition. Clustering is performed only on the initial and final segments, based on the assumption that the beginning and ending of a composition are sufficient to represent its musical nuance. Using three segments could lead to a higher number of classes, making classification more complex. Through trial-and-error experiments, the best segmentation strategy was found to be one bar for the initial segment and one bar for the final segment. The initial and final segment datasets each contain 49 pieces. Given a set X, which consists of segmented composition data, where I represents the initial segment data and F represents the final segment data, then:

$$X = (I, F)$$

$$I = (i_1, i_2, i1_3, \ldots, i_n)$$

$$F = (f_1, f_2, f_3, \ldots, f_n) \tag{1}$$

The following is an example of data segmentation results for the composition titled "Biwadha Praja," which was previously used as an example. The composition data consists of the sequence: (2, 2, 0, 0, 2, 2, 0, 3, 5, 5, 6, 5, 3, 5, 6, 7, 0, 7, 6, 5, 3, 5, 7, 6, 7, 5, 6, 7, 6, 5, 3, 2, 3, 2, 7, 6, 5, 6, 7, 2, 3, 3, 2, 7, 6, 5, 3, 2). The segmentation results for this composition are structured into three segments:

Composition data

(2, 2, 0, 0, 2, 2, 0, 3, 5, 5, 6, 5, 3, 5, 6, 7, 0, 7, 6, 5, 3, 5, 7, 6, 7, 5, 6, 7, 6, 5, 3, 2, 3, 2, 7, 6, 5, 6, 7, 2, 3, 3, 2, 7, 6, 5, 3, 2).

Composition data segmentation results:

((initial segment), (body segment), (final segment)):

((2, 2, 0, 0), (2, 2, 0, 3, 5, 5, 6, 5, 3, 5, 6, 7, 0, 7, 6, 5, 3, 5, 7, 6, 7, 5, 6, 7, 6, 5, 3, 2, 3, 2, 7, 6, 5, 6, 7, 2, 3, 3, 2, 7), (6, 5, 3, 2)).

The segmentation process was applied to all 49 compositions in the dataset. Table II presents an example of the segmentation results, displaying the initial and final segments for each composition.

TABLE II. EXAMPLE OF DATA SEGMENTATION RESULTS

| ID | Initial Segment | | | | Final Segment | | | |
|---|---|---|---|---|---|---|---|---|
| | I1 | I2 | I3 | I4 | F1 | F2 | F3 | F4 |
| G01 | 3 | 5 | 6 | 7 | 0 | 7 | 0 | 6 |
| G02 | 2 | 7 | 2 | 6 | 3 | 2 | 7 | 6 |
| G03 | 2 | 2 | 0 | 0 | 6 | 5 | 3 | 2 |
| G04 | 6 | 6 | 0 | 0 | 0 | 7 | 5 | 6 |
| G05 | 0 | 7 | 3 | 2 | 2 | 7 | 5 | 6 |
| G06 | 0 | 0 | 6 | 0 | 2 | 7 | 5 | 6 |
| G07 | 3 | 2 | 7 | 6 | 7 | 6 | 3 | 2 |
| … | … | … | … | … | … | … | … | … |
| G47 | 7 | 6 | 3 | 2 | 2 | 7 | 5 | 6 |
| G48 | 7 | 7 | 0 | 0 | 7 | 3 | 7 | 2 |
| G49 | 0 | 0 | 6 | 0 | 0 | 7 | 5 | 6 |

After obtaining the same element length in each data, data normalization was performed by removing duplicate data. Out of the 49 data points in each segment, there were 21 and 30 duplicate data points in the initial and final segments, respectively. Consequently, data normalization resulted in 28 and 19 unique data points in the initial and final segments, respectively. The initial segment consists of unique data: (G01, G02, G03, …, G10, G12, …, G16, G19, G21, …, G26, G28, G35, G38, G40, G45, G47), while the final segment consists of unique data: (G11, G17, G18, G20, G27, G29, G30, …, G34, G36, G37, G39, G41, …, G44, G46, G48, G49).

### *B. Clustering-Based Classification*

Clustering-based classification was performed using the K-Means algorithm on the initial and final segments separately, with the following steps: 1) Clustering was applied to the initial segment dataset; 2) Clustering was applied to the final segment dataset; and 3) Data concatenation was performed on the clustering output from the initial and final segments to determine the musical nuance class. The K-Means algorithm, which is a clustering method, works by grouping a set of data based on feature similarities. The features in the initial and final segments were the first and last bars in the composition, respectively, where each bar consisted of four notes. The first step in the K-Means algorithm was to initialize the centroids, which represent the mean or median of all points in the cluster. The initial segment I and final segment F contained bar data used as feature vectors. Since each bar contained four notes, each feature vector had a four-dimensional representation, which was then grouped into a number of clusters, as in the following example: I = ((3, 5, 6, 7), (2, 7, 2, 6), (2, 2, 0, 0), …, (0, 0, 6, 0)), and F = ((0, 5, 0, 2), (0, 7, 0, 6), (0, 7, 5, 6), ..., (7, 6, 7, 2)), where I and F contained 28 and 19 feature vectors, respectively. Next, the distance of each data point Xn, with X representing I and F, was calculated from each centroid using the Euclidean distance to assign the data point to the cluster with the smallest distance. The formula used is:

$$d(X_n, C_k) = \|X_n - C_k\| 2 = \sqrt{\sum_{j=1}^{d}(X_{nj} - C_{kj})^2} \quad (2)$$

where d is the feature vector dimension, k is the number of clusters, $X_{nj}$ is the j-th feature coordinate of data $X_n$, and $C_{kj}$ is the j-th feature coordinate of centroid $C_k$.

The centroid is recalculated as the average of all points in the cluster using the following formula:

$$C_k = \frac{1}{|S_k|}\sum_{X \in S_k} X_n \quad (3)$$

Where $S_k$ represents the set of data points assigned to cluster K, and $|S_k|$ is the number of elements in that cluster.

The Elbow Method is used to measure the total intra-cluster variance, also known as the Sum of Squared Errors (SSE) or Inertia. The formula for calculating total SSE or Inertia, where a smaller value of J indicates better clustering performance, is as follows:

$$J = \sum_{k=1}^{K}\sum_{X_n \in S_k}\|X_n - C_k\|^2 \quad (4)$$

The iteration process continues until the centroid remains unchanged or the distance change is minimal:

$$\left\|C_k^{(t+1)} - C_k^{(t)}\right\| < \epsilon \quad (5)$$

Where $C_k^{(t)}$ is the centroid at iteration t, and ϵ is a very small tolerance value.

The performance evaluation of the K-Means algorithm is conducted using the Silhouette Score, which measures how well a data point fits within its assigned cluster compared to other clusters. The Silhouette Score is calculated for each data point $X_n$ in the dataset using the average distance to all points within the same cluster (a) and the average distance to all points in the nearest cluster (b), as follows:

$$a(i) = \frac{1}{|C| - 1}\sum_{X_j \in C, j \neq i} d(X_n, X_j)$$

$$b(i) = \min_{C' \neq C} \frac{1}{|C'|}\sum_{X_j \in C', j \neq i} d(X_n, X_j) \quad (6)$$

where C is the cluster of $X_n$, $d(X_n, X_j)$ is the Euclidean distance between $X_n$ dan $X_j$, and $C'$ is the nearest other cluster.

The smaller the value of a(i) and the larger the value of b(i), the better the clustering results. The Silhouette Score ranges from -1 to 1, where 1 indicates well-defined clusters with data points being far from other clusters and close to their own cluster, 0 indicates that data points are on the boundary between two clusters, and -1 represents poor clustering results.

The parameter settings for the K-Means algorithm were uniformly applied to both the initial and final segments. After performing clustering, the output clusters from the initial and final segments were used as references to assign class labels to all 49 compositions based on their segments. The musical nuance class is determined by concatenating the initial segment output cluster with the final segment output cluster. For example, clustering the data for composition G002 resulted in an output cluster of 1 for the initial segment and an output cluster of 3 for the final segment. Data concatenation was performed by converting the numerical cluster output into a string format. Thus, the classification result for G002 was the musical nuance class 13. In summary, the musical nuance class is formed based on the combination of the number of clusters in the initial and final segments.

## C. Evaluations

The model uses the rule (X, Y → Z), where X and Y represent the clusters from the initial and final segments, respectively, and Z represents the class resulting from the data concatenation of X and Y. Thus, model evaluation is conducted by calculating the similarity within class Z using Euclidean distance. The smaller the average Euclidean distance within a class, the more similar the data points in that class. Given a class C with m vectors, the similarity measurement within the class is calculated using the average Euclidean distance within the class using the following formula:

$$C = (m_1, m_2, m_{3,} \dots, m_n)$$

$$D_c = \frac{1}{\binom{m}{2}} \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} d(X_i, X_j) \tag{7}$$

where $d(X_i, X_j)$ represents the Euclidean distance between the i-th and j-th data points, and $\binom{m}{2} = \frac{m(m-1)}{2}$ represents the number of unique pairs in the dataset.

The average Euclidean distance E is categorized into three groups as follows:

$$E = \begin{cases} < 3, & \text{very similar} \\ > 10, & \text{different} \\ else, & \text{similar} \end{cases} \tag{8}$$

## IV. Results

The K-Means clustering-based classification experiment resulted in a silhouette score of 0.34 for evaluating clustering performance in the initial segment and 0.5 in the final segment. The initial segment, consisting of 28 data points, formed four clusters, while the final segment, consisting of 19 data points, formed five clusters. The number of clusters in each segment was determined using the elbow method. In the initial segment, the elbow graph indicated that the inertia decline started to slow down from four clusters, whereas, in the final segment, the slowdown began at five clusters. Fig. 4 illustrates the elbow graphs for the initial and final segments.

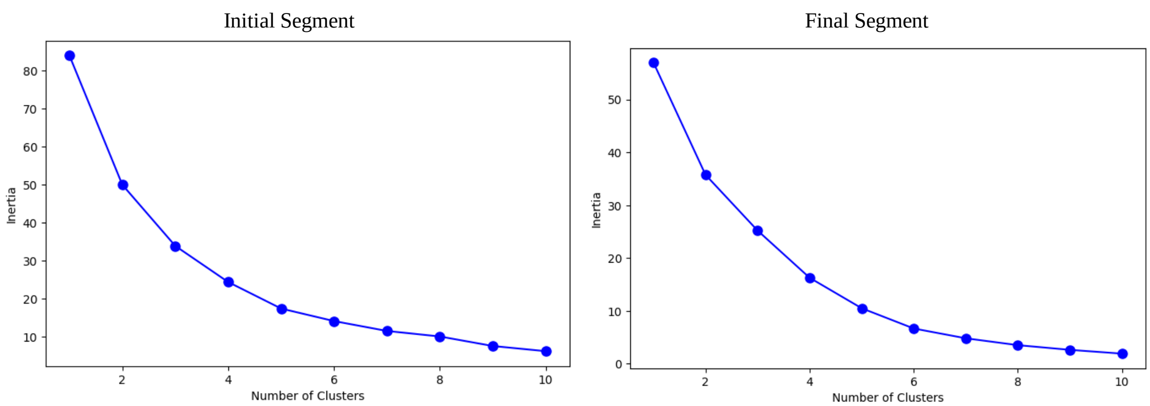


Fig. 4. Visualization of the elbow method in initial segment clustering and final segment clustering.

The initial segment, consisting of 28 data points, is distributed into clusters 0, 1, 2, and 3, with 7, 8, 6, and 7 data points in each cluster, respectively. Meanwhile, the final segment, consisting of 19 data points, is distributed into clusters 0, 1, 2, 3, and 4, with 5, 6, 2, 2, and 4 data points in each cluster, respectively. Fig. 5 illustrates the Principal Component Analysis (PCA) visualization of data distribution in both the initial and final segments.

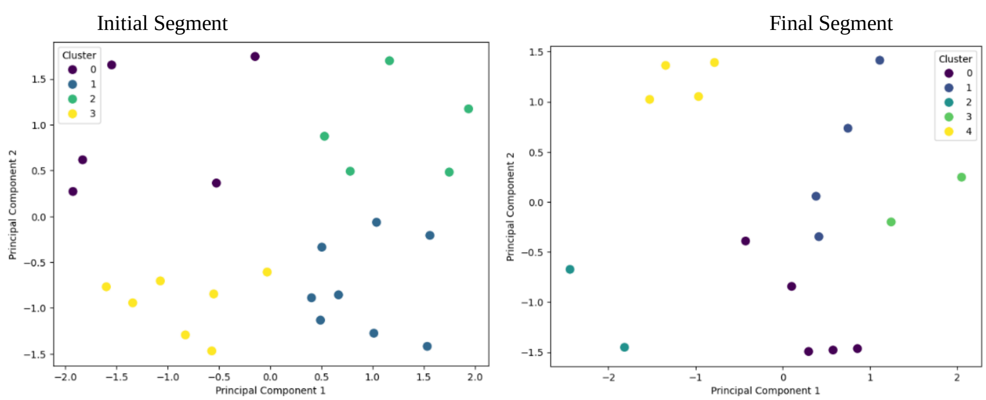


Fig. 5. Visualization of the PCA of data distribution in the initial and final segments.

Subsequently, data concatenation is applied to the cluster outputs of the initial and final segments. With four clusters in both the initial segment and the final segment, the classification results in 16 musical nuance classes derived from the string combinations of C-I and C-F, where C-I represents cluster output of the initial segmen, and C-F represents cluster output of the final segment: (00, 01, 02, 03, 10, 11, 12, 13, 20, 21, 22, 23, 30, 31, 32, 33), assuming all possible classes are formed. Each of these classes is assigned a class index: (1, 2, 3, 4, 5, …, 16). Table III illustrates the determination of musical nuance classes through data concatenation between the output clusters of the initial and final segments.

TABLE III. EXAMPLE OF DATA SEGMENTATION RESULTS

| ID | Initial Segment | | | | | Final Segment | | | | | Class Result | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | I1 | I2 | I3 | I4 | C-I | F1 | F2 | F3 | F4 | C-F | Concat. | Index |
| G01 | 3 | 5 | 6 | 7 | 2 | 0 | 7 | 0 | 6 | 2 | 22 | 11 |
| G02 | 2 | 7 | 2 | 6 | 1 | 3 | 2 | 7 | 6 | 3 | 13 | 7 |
| G03 | 2 | 2 | 0 | 0 | 3 | 6 | 5 | 3 | 2 | 1 | 31 | 14 |
| G04 | 6 | 6 | 0 | 0 | 1 | 0 | 7 | 5 | 6 | 4 | 14 | 8 |
| G05 | 0 | 7 | 3 | 2 | 3 | 2 | 7 | 5 | 6 | 4 | 34 | 18 |
| G06 | 0 | 0 | 6 | 0 | 0 | 2 | 7 | 5 | 6 | 4 | 04 | 4 |
| G07 | 3 | 2 | 7 | 6 | 0 | 7 | 6 | 3 | 2 | 1 | 01 | 2 |
| G08 | 0 | 5 | 0 | 6 | 3 | 0 | 5 | 0 | 2 | 2 | 32 | 15 |
| … | … | … | … | … | … | … | … | … | … | … | … | … |
| G47 | 7 | 6 | 3 | 2 | 1 | 2 | 7 | 5 | 6 | 4 | 14 | 8 |
| G48 | 7 | 7 | 0 | 0 | 1 | 7 | 3 | 7 | 2 | 3 | 13 | 7 |
| G49 | 0 | 0 | 6 | 0 | 0 | 0 | 7 | 5 | 6 | 4 | 04 | 4 |

## V. DISCUSSION

The classification approach for gamelan musical nuances uses initial and final segmentation to analyze melodies, followed by data concatenation of the clustering results from both segments to generate more complex labels. The dataset containing 49 gamelan music score data produced 28 unique data in the initial segment, and 19 unique data in the final segment. The high number of duplicate data points in each segment demonstrates that a high level of similarity between compositions in gamelan music is a common phenomenon. Additionally, this finding strengthens the hypothesis that musical nuances can be analyzed mathematically to identify clusters in the initial and final segments. Data segmentation shows good performance in supporting the clustering process. Table IV shows a description of the data in the initial segment and final segment.

TABLE IV. EXAMPLE OF DATA SEGMENTATION RESULTS

| | I1 | I2 | I3 | I4 | F1 | F2 | F3 | F4 |
|---|---|---|---|---|---|---|---|---|
| count | 28 | 28 | 28 | 28 | 19 | 19 | 19 | 19 |
| mean | 2.96 | 4.07 | 3 | 3.25 | 3.89 | 5.05 | 4.05 | 4.11 |
| std | 2.80 | 2.55 | 2.64 | 2.68 | 2.75 | 1.72 | 2.30 | 1.94 |
| min | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 2 |
| 25% | 0 | 2 | 0 | 1.5 | 2 | 3 | 2.5 | 2 |
| 50% | 3 | 5 | 3 | 2 | 4 | 5 | 4 | 5 |
| 75% | 6 | 6 | 6 | 6 | 6.5 | 6.5 | 6 | 6 |
| max | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 |

Analysis of the initial segment data shows that I1 has a fairly even distribution, but many small values (25% of the data have a value of 0). Most of the data fall between 0 and 6, with an average of around 3. Meanwhile, I2 tends to have higher values compared to I1, with the data centered around 4-5. There are some small values, but most of the data fall within the mid-to-high range. I3 follows a similar pattern to I1, with many small values and some high values. The data are quite spread out, with the majority ranging between 0 and 6. Furthermore, I4 has many low values but also several high values. The median of 2 indicates a tendency toward lower values, but the distribution remains broad. Overall, I1 and I3 share similar characteristics, with an average of around 3, many small values, and some high values. I2 has the highest average value (4.07), indicating a generally higher tendency compared to the other features. I4 has the lowest median (2) but remains widely distributed. This description suggests that the data are suitable for clustering. With a broad data range (0-7), there is a possibility of distinct groups. The relatively high standard deviation indicates that the data are not too homogeneous, and the varied distribution suggests the potential for meaningful patterns to be detected by a clustering algorithm.

Analysis of the final segment data shows that F1 has a fairly even distribution with an average value of 3.89. The data spread is quite wide (standard deviation 2.75), with many low values (25% of the data is below 2) but also some high values reaching 7. Meanwhile, F2 tends to have higher values compared to F1, with an average of 5.05 and a lower standard deviation (1.72), indicating that the data is more concentrated around the central value (median 5). F3, similar to F1, with an average of 4.05 and a standard deviation of 2.30, indicating a fairly wide spread. The data ranges from 0 to 7, with many low values (Q1 = 2.5) but also a significant number of high values. Furthermore, F4 has an average of 4.11, slightly higher than F3, with a standard deviation of 1.94. The data distribution is more concentrated compared to F1 and F3 but still shows considerable variation. Overall, F2 has the highest average value (5.05) and the smallest standard deviation, indicating a more centralized distribution, while F1 and F3 share similar patterns, with a wider spread and many low values, and F4 falls between F2 and F3, with a tendency to be more concentrated but still showing a fairly wide distribution. The data is quite varied, with a wide range of values

(0-7) in all features, suggesting the potential for distinct patterns to be identified through clustering.

TABLE V. Z-SCORE NORMALITATION RESULTS IN INITIAL SEGMENT

| NO | ID | I1 | I2 | I3 | I4 |
|---|---|---|---|---|---|
| 1 | G01 | 0.01 | 0.37 | 1.16 | 1.43 |
| 2 | G02 | -0.35 | 1.17 | -0.39 | 1.05 |
| 3 | G03 | -0.35 | -0.83 | -1.16 | -1.24 |
| 4 | G04 | 1.11 | 0.77 | -1.16 | -1.24 |
| 5 | G05 | -1.08 | 1.17 | 0.00 | -0.48 |
| … | … | … | … | … | … |
| 26 | G40 | -1.08 | -0.43 | -1.16 | -0.48 |
| 27 | G45 | 1.47 | -0.43 | 1.54 | -0.48 |
| 28 | G47 | 1.47 | 0.77 | 0.00 | -0.48 |

TABLE VI. Z-SCORE NORMALITATION RESULTS IN FINAL SEGMENT

| No | ID | F1 | F2 | F3 | F4 |
|---|---|---|---|---|---|
| 1 | G01 | -1.46 | 1.17 | -1.81 | 1.00 |
| 2 | G02 | -0.33 | -1.83 | 1.32 | 1.00 |
| 3 | G03 | 0.79 | -0.03 | -0.47 | -1.11 |
| 4 | G04 | -1.46 | 1.17 | 0.42 | 1.00 |
| 5 | G05 | -0.71 | 1.17 | 0.42 | 1.00 |
| … | … | … | … | … | … |
| 17 | G40 | -1.08 | -0.43 | -1.16 | -0.48 |
| 18 | G45 | 1.47 | -0.43 | 1.54 | -0.48 |
| 19 | G47 | 1.47 | 0.77 | 0.00 | -0.48 |

Data standardization was carried out using Z-score normalization with a standard scale with an average value of 0 and a standard deviation of 1. Negative values indicate that the data is below average, while positive values are above average. Table V and Table VI shows Z-score normalization results in the initial segments and final segments, respectively. Next, cluster determination for each data is done by calculating the closest Euclidean distance to each cluster. Table VII and Table VIII show examples of cluster determination results for each data in the initial segment and final segment, respectively.

TABLE VII. Z-SCORE NORMALITATION RESULTS IN INITIAL SEGMENT

| NO | ID | 0 | 1 | 2 | 3 | Cluster |
|---|---|---|---|---|---|---|
| 1 | G01 | 2.04 | 1.88 | 0.94 | 2.34 | 2 |
| 2 | G02 | 2.74 | 1.11 | 2.19 | 1.42 | 1 |
| 3 | G03 | 1.80 | 2.05 | 3.09 | 0.94 | 3 |
| 4 | G04 | 3.31 | 0.75 | 2.46 | 2.09 | 1 |
| 5 | G05 | 2.63 | 1.89 | 2.46 | 1.58 | 3 |
| … | … | … | … | … | … | … |
| 26 | G40 | 1.94 | 2.26 | 3.31 | 0.50 | 3 |
| 27 | G45 | 2.65 | 2.54 | 1.17 | 3.44 | 2 |
| 28 | G47 | 3.15 | 0.94 | 1.41 | 2.62 | 1 |

TABLE VIII. Z-SCORE NORMALITATION RESULTS IN FINAL SEGMENT

| NO | ID | 0 | 1 | 2 | 3 | 4 | Cluster |
|---|---|---|---|---|---|---|---|
| 1 | G01 | 2.87 | 3.35 | 0.60 | 4.54 | 2.49 | 2 |
| 2 | G02 | 2.17 | 2.79 | 4.10 | 0.81 | 3.16 | 3 |
| 3 | G03 | 1.23 | 0.88 | 2.68 | 2.36 | 2.49 | 1 |
| 4 | G04 | 2.84 | 2.62 | 2.32 | 3.40 | 0.44 | 4 |
| 5 | G05 | 2.52 | 1.93 | 2.43 | 3.05 | 0.44 | 4 |
| … | … | … | … | … | … | | |
| 17 | G45 | 2.75 | 1.07 | 4.08 | 2.23 | 2.42 | 1 |
| 18 | G46 | 1.23 | 0.88 | 2.68 | 2.36 | 2.49 | 1 |
| 19 | G48 | 2.28 | 1.91 | 4.46 | 0.81 | 3.35 | 3 |

There are four clusters in the initial segment and five clusters in the final segment, resulting in a maximum of 20 possible classes from the data concatenation of both segments. However, data concatenation on clusters in the initial and final segments produces 17 classes from a possible 20 clusters. The 17 clusters formed are: (00, 01, 02, 04, 10, 11, 13, 14, 20, 21, 22, 24, 30, 31, 32, 33, 34), and are given class indices: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, and 17. Three classes: Class 4, Class 8, and Class 14, were not formed. These classes would have been composed of cluster 0, 1 and 2 in the initial segment, and cluster 3, 2, and 3 in the final segment, respectively. The absence of these classes could be due to factors such as the lack of melodic variations corresponding to these clusters or the limited dataset size, which may not cover all cluster combinations.

The distribution of data from 49 compositions into classes ranges from 1 to 8 data points per class. Class 5 contains the most data, while Class 2, Class 13, Class 16, and Class 19 each contain only one data point. Classes with only one data point cannot be evaluated for similarity, which may indicate that certain cluster combinations are rare or underrepresented in the dataset. The class similarity evaluation using the average Euclidean distance produced fairly good results. An average Euclidean distance below three indicates a high level of similarity, while a value greater than ten signifies significant differences. Three of the 13 formed classes with more than one data point, Class 5, Class 12, and Class 18, demonstrate high similarity with a value below three. Five other classes, Class 7, Class 10, Class 11, Class 15, and Class 20, exhibit near-high similarity, with values ranging from three to four, while the remaining five classes fall within the range of four to five. Table IX shows the classification results based on data concatenation, where N represents the number of melodies, and E represents the average Euclidean distance.

The clustering results in the initial and final segments show that the same musical nuance can be represented through both identical and different notation sequences. Data concatenation produces more complex and diverse classes, where a musical nuance that begins with the same characteristics in the initial segment can shift in the final segment. For example, out of 49 melodies, in the initial segment, cluster 0 is represented by 15 measures, eight of which are unique measures: (0, 0, 2, 7), (0, 0, 3, 2), (0, 0, 3, 5), (0, 0, 6, 0), (0, 0, 6, 5), (0, 0, 2, 7), (3, 2, 7, 6), and (3, 2, 3, 7). Furthermore, the 15 measures in the final segment, which are their counterparts, belong to clusters 0, 1, 2, and 4. Thus, these 15 melodies form four musical nuance classes: Class 1 (00), Class 2 (01), Class 3 (02), Class 5 (04). Notably, Class 4 (03) does not exist, which would have been formed by cluster 0 in the initial segment and cluster 3 in the final segment. This indicates that musical nuances in cluster 0 of the initial segment can transition into any cluster in the final segment except cluster 3. A similar pattern is observed in melodies where the initial segment belongs to cluster 1 and cluster 4. In the initial segment, only cluster 3 is flexible enough to be paired with all clusters in the final segment. Out of 49 melodies, 11 have measures in the initial segment that belong to cluster 3. These 11 measures consist of five unique notation sequences: (0, 2, 0, 7), (0, 3, 0, 2), (0, 5, 0, 6), (0, 6, 0, 7), (0, 7, 3, 2), and (3, 3, 0, 0).

TABLE IX. EXAMPLE OF DATA SEGMENTATION RESULTS

| Class Results. | | N | E |
|---|---|---|---|
| **Concat.** | **Index** | | |
| 00 | Class 1 | 4 | 4.6 |
| 01 | Class 2 | 1 | - |
| 02 | Class 3 | 3 | 4.9 |
| 03 | Class 4 | 0 | - |
| 04 | Class 5 | 8 | 2.2 |
| 10 | Class 6 | 3 | 4.3 |
| 11 | Class 7 | 5 | 3.4 |
| 12 | Class 8 | 0 | - |
| 13 | Class 9 | 2 | 4.9 |
| 14 | Class 10 | 6 | 3.0 |
| 20 | Class 11 | 2 | 3.3 |
| 21 | Class 12 | 2 | 2.3 |
| 22 | Class 13 | 1 | - |
| 23 | Class 14 | 0 | - |
| 24 | Class 15 | 2 | 3.5 |
| 30 | Class 16 | 1 | - |
| 31 | Class 17 | 2 | 4.9 |
| 32 | Class 18 | 3 | 2.4 |
| 33 | Class 19 | 1 | - |
| 34 | Class 20 | 3 | 3.3 |
| **Sum** | | **49** | |

Below is an illustration of the classification results for musical nuances based on segmentation and data concatenation, presented in the C: I $\rightarrow$ F format, where C represents the class (musical nuance in the melody), I represents the notation sequence in the initial segment, and F represents the notation sequence in the final segment.

Class 16: (3, 3, 0, 0) $\rightarrow$ (4, 3, 2, 7)

Class 19: (3, 3, 0, 0) $\rightarrow$ (3, 2, 7, 6)

Class 20: (3, 3, 0, 0) $\rightarrow$ (2, 7, 5, 6)

Class 20: (3, 3, 0, 0) $\rightarrow$ (2, 7, 5, 6)

Class 17: (0, 2, 0, 7) $\rightarrow$ (0, 7, 0, 6)

Class 17: (0, 3, 0, 2) $\rightarrow$ (0, 7, 0, 6)

Class 17: (0, 5, 0, 6) $\rightarrow$ (0, 5, 0, 2)

## VI. CONCLUSION

A clustering-based classification model was developed in this study. A novel approach was proposed by incorporating data segmentation and concatenation, which proved effective in controlling gamelan musical nuance classification. The clustering process using the K-Means algorithm demonstrated good performance and results, while the classification rules derived from data segmentation and concatenation helped reduce clustering complexity, yielding promising classification outcomes.

Overall, the data segmentation and concatenation model in K-Means clustering-based gamelan musical nuance classification shows promising results. Some issues, such as the absence of certain classes due to missing cluster combinations between the initial and final segments, as well as the presence of classes with only one data point, may stem from the relatively small dataset of 49 compositions. Collecting symbolic data for gamelan music remains a challenge. Unlike Western music, where musical data is well-managed and documented, with easy access to public datasets, gamelan music data for research purposes is not yet well-administered, making information access limited. This condition may also apply to other traditional music, such as Chinese traditional music [25]. However, the data segmentation and concatenation approach for controlling clustering-based musical emotion classification has the potential to enrich analysis by considering the relationship between the initial and final segments in a composition.

For future work, this approach can be applied to larger music datasets while maintaining the same segmentation and clustering techniques. Additionally, the proposed method provides positive opportunities to be implemented for other types of music, including Western music. The proposed method can also enhance the implementation of style imitation techniques in automatic music generation by controlling musical emotions when selecting melodies as the dataset, where the dataset consists of a collection of melodies sourced from the same composer.

## REFERENCES

[1] Y. Xia, and F. Xu, “Study on Music Emotion Recognition Based on the Machine Learning Model Clustering Algorithm, Mathematical Problems,” Engineering, vol. 2022, 9256586, 2022. Doi: 10.1155/2022/9256586.

[2] Y.O. Medina, J. R. Beltrán, and S. Baldassarri, “Emotional classification of music using neural networks with the MediaEval dataset,” Pers Ubiquit Comput., vol. 26, pp. 1237–1249, 2022. Doi: 10.1007/s00779-020-01393-4.

[3] X. Jia, “Music Emotion Classification Method Based on Deep Learning and Improved Attention Mechanism,” Computational Intelligence and Neuroscience, vol. 2022, 5181899, 2022. Doi: 10.1155/2022/5181899.

[4] D. Han, Y. Kong, J. Han, and G. Wang, “A survey of music emotion recognition,” Front. Comput. Sci., vol. 16, 166335, 2022. Doi: 10.1007/s11704-021-0569-4.

[5] E. Parada-Cabaleiro, A. Batliner, M. Zentner, and M. Schedl, “Exploring emotions in Bach chorales: a multi-modal perceptual and data-driven

study," Royal Society Open Science, vol. 10, no. 12, 230574, 2023. Doi: 10.1098/rsos.230574

[6] C. Bao, and Q. Sun, "Generating Music With Emotions," in IEEE Transactions on Multimedia, vol. 25, pp. 3602-3614, 2023, Doi: 10.1109/TMM.2022.3163543.

[7] S. Sulun, M. E. P. Davies, and P. Viana, "Symbolic Music Generation Conditioned on Continuous-Valued Emotions," in IEEE Access, vol. 10, pp. 44617-44626, 2022, doi: 10.1109/ACCESS.2022.3169744.

[8] D. Chaudhary, N. P. Singh, and S. Singh, "Development of music emotion classification system using convolution neural network," Int J Speech Technol., vol. 24, pp. 571–580, 2021. Doi: 10.1007/s10772-020-09781-0.

[9] Y. Huo, "Music Personalized Label Clustering and Recommendation Visualization," Complexity, vol. 2021, 5513355, 2021. Doi: 10.1155/2021/5513355.

[10] S. Nag, M. Basu, S. Sanyal, A. Banerjee, and D. Ghosh, "On the application of deep learning and multifractal techniques to classify emotions and instruments using Indian Classical Music," Physica A: Statistical Mechanics and its Applications, vol. 597, 127261, 2022. Doi: 10.1016/j.physa.2022.127261.

[11] A. Irama, and M. Form, "Temporal and Density Flow in Javanese Gamelan," Available at https://sumarsam.faculty.wesleyan.edu/files/2023/01/4_Temporal_and_Density_Flow.pdf.

[12] M. Benamou, "Rasa: affect and intuition in Javanese musical aesthetics," Oxford University Press, 2010.

[13] S. Suyoto, and A. Setiawan, "The Meaning of Gendhing Kodhok Ngorek in the Panggih Procession of a Traditional Javanese Wedding Ceremony," Journal of Urban Society's Arts, vol. 10, no. 1, pp. 53-62, 2023.

[14] J. Qiu, C. L. Chen, and T. Zhang, "A novel multi-task learning method for symbolic music emotion recognition," arXiv preprint arXiv:2201.05782., 2022.

[15] J. Lian, "An artificial intelligence-based classifier for musical emotion expression in media education," PeerJ Computer Science, vol. 9, e1472, 2023. Doi: 10.7717/peerj-cs.1472 .

[16] A. Z. Fanani, A. M. Syarif, G. F., and A. Marjuni, "Expressing and Developing Melodic Phrases in Gamelan Skeletal Melody Generation Using Genetic Algorithm," IEEE Access, vol. 12, pp. 130512-130523, 2024. Doi: 10.1109/ACCESS.2024.3457880.

[17] A. M. Syarif, A. Azhari, S. Suprapto, and K. Hastuti, "Gamelan Melody Generation Using LSTM Networks Controlled by Composition Meter Rules and Special Notes," Journal of Advances in Information Technology, vol. 14, no. 1, pp. 26-38, 2023. Doi: 10.12720/jait.14.1.26-38.

[18] S. Ndhlovu, and R. Ajoodha, "A Novel Feature-Based Music Recommendation System Considering the Uniqueness of Musical Items," 2022. Available at SSRN: https://ssrn.com/abstract=4332853. Doi: 10.2139/ssrn.4332853.

[19] N. Ferreira, L., Mou, L., Whitehead, J., and L. H. S., Lelis, L. H. S, "Controlling Perceived Emotion in Symbolic Music Generation with Monte Carlo Tree Search," in Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, vol. 18, no. 1, pp. 163-170, 2022. Doi: 10.1609/aiide.v18i1.21960.

[20] E. Cambria, X. Zhang, R. Mao, M. Chen, and K. Kwok, "SenticNet 8: Fusing emotion AI and commonsense AI for interpretable, trustworthy, and explainable affective computing," in International Conference on Human-Computer Interaction, pp. 197-216, 2022.

[21] L. Ma, W. Zhong, X. Ma, L. Ye, and Q. Zhang, "Learning to generate emotional music correlated with music structure features," Cognitive Computation and Systems, vol. 4, no. 2, pp. 100-107, 2022. Doi: 10.1049/ccs2.12037.

[22] H. Jeong, H., and J. H. Yoo, J. H, "Opera Clustering: K-means on librettos datasets," Journal of Internet Computing and Services, vol. 23, no. 2, pp. 45–52, 2022. Doi: 10.7472/JKSII.2022.23.2.45.

[23] Y. Jiang, Y., and X. Jin, "Using k-Means Clustering to Classify Protest Songs Based on Conceptual and Descriptive Audio Features," in: Rauterberg, M. (eds) Culture and Computing. HCII 2022. Lecture Notes in Computer Science, vol 13324, 2022, Springer, Cham. Doi: 10.1007/978-3-031-05434-1_19.

[24] J. Sun, "Personalized music recommendation algorithm based on spark platform," Computational Intelligence and Neuroscience, vol. 2022, no. 1, 7157075. 2022. Doi: 10.1155/2022/7157075.

[25] D. Wu, X. Jia, W. Rao, W. Dou, Y. Li, and B. Li, "Construction of a Chinese traditional instrumental music dataset: A validated set of naturalistic affective music excerpts," Behavior Research Methods, pp. 1-22. 2024. Doi: 10.3758/s13428-024-02411-6.