Indonesian Journal of Data and Science

Volume 6 Issue 3
ISSN 2715-9936
https://doi.org/10.56705/ijodas.v6i3.358

Research Article

Classification of Gamelan Selonding Music Using Convolutional Neural Network

Ni Putu Diah Pradnya Savitri [1,*]; Anak Agung Gde Bagus Ariana [2]; Ni Kadek Nita Noviani Pande [3]; I Made Dwi Putra Asana [4]; I Gusti Agung Indrawan [5]

[1] Institut Bisnis Dan Teknologi Indonesia, Denpasar, Indonesia, diahpradnyasavit27@gmail.com
[2] Institut Bisnis Dan Teknologi Indonesia, Denpasar, Indonesia, gungariana@instiki.ac.id
[3] Institut Bisnis Dan Teknologi Indonesia, Denpasar, Indonesia, novy.pande@instiki.ac.id
[4] Institut Bisnis Dan Teknologi Indonesia, Denpasar, Indonesia, dwiputraasana@instiki.ac.id
[5] Institut Bisnis Dan Teknologi Indonesia, Denpasar, Indonesia, agung.indrawan@instiki.ac.id

Correspondence should be addressed to Ni Putu Diah Pradnya Savitri; diahpradnyasavit27@gmail.com

Received 10 August 2025; Accepted 15 October 2025; Published 31 December 2025

Abstract:
Introduction: Balinese Selonding gamelan is an endangered sacred repertoire, and automatic recognition of its musical pieces can support documentation and preservation. Method: This study investigates the automatic classification of Selonding gamelan music using a Convolutional Neural Network (CNN). The dataset consists of 10 traditional Selonding compositions. Recordings were segmented into fixed 15-second excerpts, converted to WAV, normalized, and transformed into time–frequency features using two approaches: Mel-Frequency Cepstral Coefficients (MFCC) and Constant-Q Transform (CQT). A CNN-based classifier was trained and evaluated using 5-fold cross-validation for each feature representation. Results: The MFCC-based model achieved stable high performance, with mean accuracy of 94.67% (±2.11%), mean precision of 94.97% (±1.90%), mean recall of 94.67% (±2.11%), and mean F1-score of 94.63% (±2.12%) across folds. In contrast, the CQT-based model performed notably worse, reaching only 58.04% mean accuracy and 53.28% mean F1-score, with large variance across folds. These results indicate that MFCC features capture the discriminative timbral characteristics of Selonding more effectively than CQT under the current experimental setting. Conclusion: Overall, the findings show that a CNN trained on MFCC features can reliably distinguish Selonding compositions using only short (15-second) audio segments, despite limited data. This suggests that deep learning is a feasible strategy for indexing, retrieval, and long-term preservation of Balinese gamelan repertoires.

Keywords: Gamelan Selonding, CNN, MFCC, CQT, Audio Classification, Cultural Preservation

Dataset link: https://drive.google.com/drive/folders/1Ci0hBqgZu-L3reAhytCowhyUVDudbB9R?usp=drive_link y

1. Introduction

Bali possesses a rich and diverse cultural heritage, one of which is Balinese gamelan music, a traditional karawitan orchestra tuned in pelog and slendro scales, serving as entertainment, ceremonial accompaniment, and a symbol of identity and spirituality that continues to be preserved [1], [2], [3]. Gamelan is a traditional Balinese musical ensemble that produces distinctive harmonies through the combination of various instruments with unique tones and playing techniques [4], [5]. One of the oldest and most sacred Balinese gamelan ensembles is Gamelan Selonding, originating from Tenganan Pegringsingan Village, Karangasem Regency, Bali, a Bali Aga community consisting only of native inhabitants who have maintained pre-Hindu traditions passed down through generations [6], [7]. Gamelan Selonding is considered the oldest form of traditional Balinese music, made of iron with a seven-tone pelog scale. It possesses distinctive musical characteristics, is sanctified by the Bali Aga community, and functions as a sacred offering medium, ritual accompaniment, and traditional dance music, while being revered as a symbol of spirituality and ancestral heritage [8], [9], [10].

However, the preservation of Gamelan Selonding faces challenges due to declining interest among younger generations, a limited number of performers, minimal documentation, and difficulties in musical classification. Meanwhile, the influence of modern technology and globalization underscores the importance of maintaining the relevance and protection of traditional arts [11], [12], [13]. The digitalization and analysis of gamelan, which carries unique characteristics and profound social, educational, and cultural values, represent crucial steps toward preserving Balinese traditional music heritage. Therefore, an automatic classification system is needed to identify and categorize Gamelan Selonding music as a tool for cultural learning and preservation [14], [15], [16]. On the other hand, recent developments in artificial intelligence (AI) technology, particularly in the fields of pattern recognition and audio processing, offer innovative solutions for music classification through the application of Convolutional Neural Networks (CNNs), which are capable of extracting and recognizing the unique characteristics of sound signals [17]. CNN is a deep learning method designed to automatically extract and identify complex patterns through convolutional layers, making it effective for processing both image and audio data [18], [19], [20].

Numerous studies have demonstrated that the Convolutional Neural Network (CNN) method is effective for complex audio pattern recognition and classification, applicable to both traditional and modern music. CNNs can accurately extract spectral features and recognize distinctive patterns in various regional and genre [21], [22], [23]. Research also shows that CNNs using MFCC features are effective in classifying diverse types of audio, ranging from wind instruments, baby cries, and owl sounds, to environmental audio and piano chord detection [24], [25], [26], [27], [28]. The integration of Convolutional Neural Networks (CNNs) with MFCC and CQT features has proven effective in representing timbre, pitch, and complex frequency patterns [29], [30], [31]. Nevertheless, to date, no study has specifically examined the classification of Gamelan Selonding compositions in the Tenganan style, which possess unique and distinctive musical characteristics compared to other gamelan ensembles. Based on the aforementioned background, this research aims to examine the application of Convolutional Neural Networks (CNNs) in classifying Gamelan Selonding compositions in the Tenganan style.

The main problem addressed in this research is the absence of an automatic system capable of accurately classifying Gamelan Selonding music titles. The rhythmic complexity, similarity between compositions, and limited availability of audio data present significant challenges in the preservation and digitalization of this traditional music.

The primary objective of this research is to design and implement a Gamelan Selonding music classification system using the Convolutional Neural Network (CNN) method. Specifically, the study aims to apply deep learning techniques to analyze the audio patterns of Gamelan Selonding through MFCC and CQT representations, and to evaluate the performance of the CNN model in recognizing and distinguishing between Gamelan Selonding compositions in the Tenganan style.

This study seeks to address two main research questions related to the application of deep learning technology to Balinese traditional music. First, how can the Convolutional Neural Network (CNN) method be applied to classify Gamelan Selonding music based on its audio characteristics, such as frequency patterns, timbre, and distinctive sound dynamics. Second, to what extent can the CNN model accurately differentiate between the titles of Gamelan Selonding compositions in the Tenganan style, thereby evaluating the effectiveness of this approach in recognizing musical variations within this traditional ensemble

The scope of this research focuses on the classification process of Gamelan Selonding music in the Tenganan style using audio recordings obtained from the YouTube platform. The dataset consists of ten Gamelan Selonding compositions, each segmented into 15-second audio clips, resulting in a total of 200 audio files. The analysis centers on audio feature extraction using the Mel-Frequency Cepstral Coefficients (MFCC) and Constant-Q Transform (CQT) methods to identify acoustic characteristics, without involving aspects such as musical notation, choreography, or the philosophical meaning of the music. Furthermore, this study employs only the Convolutional Neural Network (CNN) architecture as the primary model, without comparing it to other machine learning methods.

This study provides contributions in two main aspects. From an academic perspective, it expands the application of Convolutional Neural Networks (CNNs) in the domain of Balinese traditional music, particularly Gamelan Selonding, which has rarely been explored in the field of computational music. From a practical perspective, this research contributes to cultural preservation through the development of an automatic classification system that can be utilized for archiving, educational, and promotional purposes, enabling traditional music to gain wider recognition in society. Previous studies on traditional music classification using CNNs have primarily focused on Indian ragas, Chinese guqin, or Western folk datasets. However, these approaches rarely address instruments with iron timbres and cyclical rhythmic patterns such as the Gamelan Selonding of Bali. Unlike bamboo or bronze gamelan ensembles, Selonding employs iron metallophones tuned to non-equal temperaments, producing overlapping overtones and percussive envelopes that challenge feature extraction.

Therefore, this research aims to (1) evaluate the effectiveness of MFCC and CQT features in representing Selonding's acoustic characteristics, and (2) implement and assess a CNN architecture for classifying Selonding compositions from the Tenganan tradition.

2. Method:

Research design

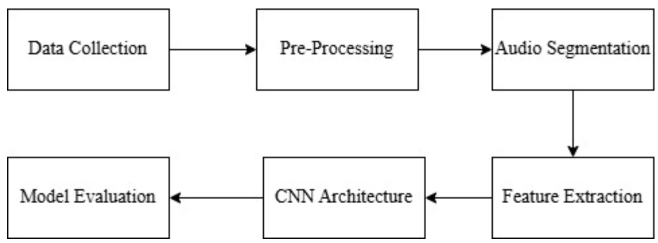

Figure 1. Research Flow

Figure 1 illustrates the research flow for Classification of Gamelan Selonding Music Using Convolutional Neural Network. The research process begins with data collection in the form 10 of Gamelan Selonding Tenganan-style music recordings obtained from YouTube. Each recording is segmented into 15. The next step involves feature extraction using the Mel-Frequency Cepstral Coefficients (MFCC) and Constant-Q Transform (CQT) methods to capture the frequency and spectral characteristics of each audio segment. These extracted features are then used as input for a CNN. The final stage includes training and testing the model to evaluate CNN performance based on accuracy, precision, and loss metrics, thereby assessing the effectiveness of the model in automatically classifying Balinese traditional music.

Sample or data selection

The dataset in this study consists of 10 audio recordings of Gamelan Selonding in the Tenganan style sourced from YouTube, selected based on recording quality, authenticity of performance style, and instrument clarity. The videos were converted to MP3 using Y2Mate and then to high-quality WAV format using Audacity to ensure lossless audio compatible with Python libraries such as Librosa. Each audio file was standardized to mono at a 22,050 Hz sampling rate, then pre-processed by amplitude-normalizing the signal (with DC-offset removal) and, when necessary, lightly trimming excessive silence or obvious clicks without altering musical content, and segmented into 20 clips of 15 seconds taken proportionally from the beginning, middle, and end of each piece to capture musical variation, producing a total of 200 audio segments for analysis.

Table 1 presents the list of the ten Selonding compositions used as the dataset, which includes the class (composition title) and total duration of each recording. From these audio segments, two types of features were extracted: Mel-Frequency Cepstral Coefficients (MFCC), to model timbral and spectral envelope characteristics relevant to human auditory perception, and Constant-Q Transform (CQT), to model harmonic structure on a logarithmic frequency scale. These feature representations were then used as input to the CNN during training and testing for automatic classification of Selonding compositions.

Table 1. Data Set

Class	Duration
Lagu001_Gending Sekar Gadung	15 Minute
Lagu002_Gending Rejang Reong	10 Minute
Lagu003_Gending Rejang Gucek	11 Minute
Lagu004_Gending Rejang Renteng	12 Minute
Lagu005_Gending Rejang Lente	10 Minute
Lagu006_Gending Kelompok Guna	9 Minute
Lagu007_Gending Nyangjangan	8 Minute
Lagu008_Gending Mekare-kare	11 Minute
Lagu009_Gending Rejang Ileh	10 Minute
Lagu010_Gending Rejang Dauh Tukad	12 Minute

Tools and technology used

Below is the list of tools and technologies used in this research. Each item reflects what was actually implemented in the experimental pipeline, including data processing, model training, evaluation, and reporting.

a) Programming language: Python (NumPy, librosa, pandas, matplotlib, seaborn).
b) Audio feature extraction: MFCC and Constant-Q Transform (CQT).
c) Model architecture: Custom Convolutional Neural Network (Conv2D → MaxPooling2D → Dropout → Dense → Softmax).
d) Activation function: ReLU in all convolutional and dense layers.
e) Optimizer: Adam with Early Stopping (restoring best weights).
f) Deep learning framework: TensorFlow and Keras (Sequential API).
g) Five-fold cross-validation with folds grouped by source recording to prevent data leakage.Performance metrics: accuracy, precision, recall, F1-score, and confusion matrix per fold.
h) Visualization tools: matplotlib and seaborn for training curves and confusion matrices.
i) Reporting workflow: automatic export of per-fold and summary results to Excel.

The Convolutional Neural Network (CNN) developed in this study was designed to classify the Selonding audio segments based on their MFCC and CQT feature representations. The model consists of three convolutional blocks, each comprising a Conv2D layer with 32, 64, and 128 filters respectively, using a 3×3 kernel size and ReLU activation to capture local time–frequency patterns within the spectrograms. Each convolutional layer is followed by a 2×2 max-pooling layer to progressively reduce spatial dimensionality, as well as a dropout layer with a rate of 0.25 to mitigate overfitting. After the convolutional blocks, the output is flattened and passed to a fully connected dense layer with 128 units and ReLU activation, accompanied by a dropout rate of 0.5 to further enhance generalization. The final output layer uses a softmax activation function with ten units corresponding to the number of Selonding composition classes. The model was trained using the Adam optimizer with a learning rate of 0.001, a batch size of 32, and up to 75 epochs, with early stopping applied (patience = 10) to prevent overfitting during training. This architecture balances representational capacity and computational efficiency, making it suitable for learning discriminative features from relatively small traditional music datasets.

Data collection process

The dataset in this study comprises ten audio recordings of Gamelan Selonding in the Tenganan style, sourced from YouTube and selected based on recording quality, authenticity of performance style, and clarity of instrumental sound. All recordings were converted from MP3 to high-quality WAV format using Audacity to preserve lossless fidelity and ensure compatibility with Python audio-processing libraries such as Librosa. The use of publicly available content was restricted to academic, non-commercial research purposes in accordance with fair use guidelines. Each recording was standardized to a mono channel with a sampling rate of 22,050 Hz and preprocessed through DC-offset removal and amplitude normalization to minimize recording inconsistencies. Subsequently, the audio was segmented into twenty clips of fifteen seconds each, proportionally sampled from the beginning, middle, and end of every composition to capture structural and timbral variations, yielding a total of 200 audio segments for analysis. No data augmentation techniques were applied at this stage. From each segment, Mel-Frequency Cepstral Coefficients (MFCC; 40 coefficients) and Constant-Q Transform (CQT; 84 frequency bins, spanning C1–C7) were extracted and converted into two-dimensional spectrogram matrices to serve as input features for the Convolutional Neural Network (CNN) model.

Data analysis methods

The data analysis methods in this study comprised several key stages, beginning with audio preprocessing, feature extraction, and continuing to CNN model training and evaluation. During the preprocessing stage, segmented audio data were normalized and converted into numerical format to enable computational processing. Feature extraction was then conducted using two main techniques: Mel-Frequency Cepstral Coefficients (MFCC) and Constant-Q Transform (CQT). MFCC was used to capture timbral and spectral characteristics of sound based on human auditory perception, while CQT was employed to represent harmonic structures and logarithmic frequency distributions corresponding to the unique tonal qualities of Gamelan Selonding.

The extracted features were converted into two-dimensional matrices, which served as inputs to the Convolutional Neural Network (CNN) model. The CNN architecture was implemented using TensorFlow and Keras libraries, consisting of multiple convolutional, pooling, and fully connected layers designed to identify acoustic patterns within each audio segment. After the training process, the model's performance was evaluated using accuracy, precision, recall, and loss metrics to measure its effectiveness in automatically classifying the titles of Gamelan Selonding compositions.

3. Result and Discussion:

Result

Description of the data processing results

The audio dataset consists of 10 traditional Selonding gamelan compositions. Each recording was segmented into fixed 15-second excerpts, converted to WAV format, normalized, and then transformed into two different time–frequency feature representations: Mel-Frequency Cepstral Coefficients (MFCC) and Constant-Q Transform (CQT). These features were used as inputs to a Convolutional Neural Network (CNN), and model performance was evaluated using 5-fold cross-validation to measure generalization stability across folds.

Visualization of the results

Figure 2(a) shows the confusion matrix of the model using MFCC features, while **Figure 2(b)** presents the confusion matrix using CQT features. The darker diagonal cells indicate higher classification accuracy for each Selonding composition, whereas the off-diagonal elements represent misclassifications between classes. Meanwhile, **Figure 3(a)** and **Figure 3(b)** display the training and validation accuracy of the CNN model using MFCC and CQT features, respectively, while **Figure 3(c)** and **Figure 3(d)** show the corresponding training and validation loss curves. Both models exhibit effective learning behavior, where the MFCC-based model demonstrates smoother convergence and higher stability, whereas the CQT-based model captures richer spectral representations with slight fluctuations during training.

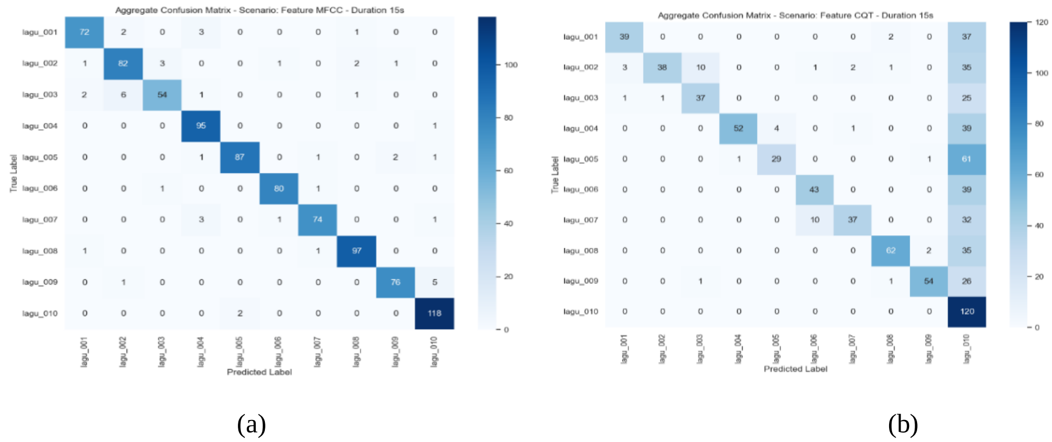

Figure 2. Confusion Matrix of the models: Fitur MFCC **(**a) and Fitur CQT (b)

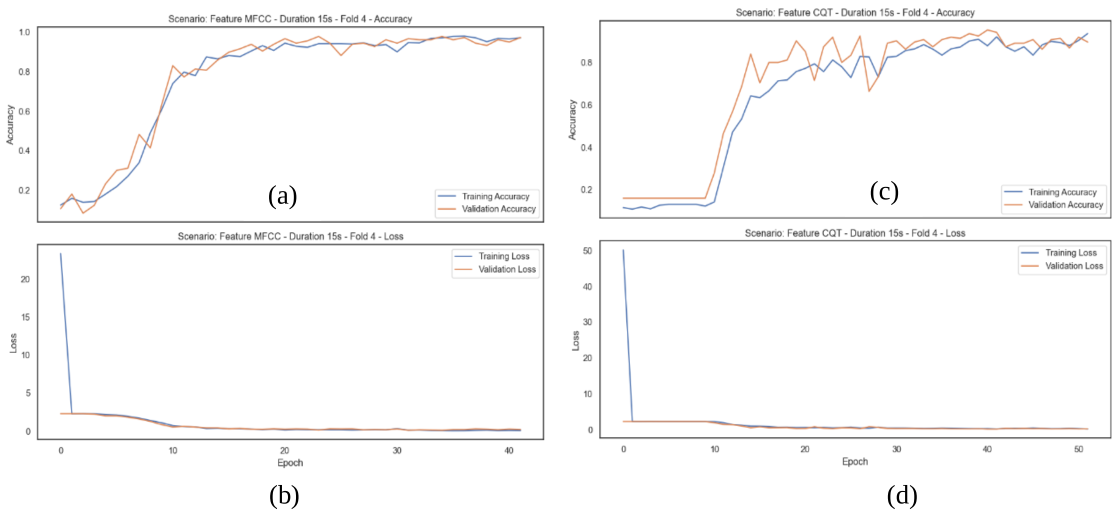

Figure 3. Training and validation performance of the CNN model. (a) Accuracy with MFCC; (b) Loss with MFCC; (c) Accuracy with CQT; (d) Loss with CQT.

Interpretation of the results

Based on the results presented in **Table 2**, the CNN trained on MFCC features demonstrated high and stable performance, achieving a mean accuracy of 94.67% (standard deviation 2.11%), mean precision of 94.97%, mean recall of 94.67%, and mean F1-score of 94.63%. As shown in **Table 3**, the per-fold accuracy of the MFCC-based model ranged from 91.48% to 96.59%, indicating that the model maintained consistent performance across different folds and was not overly sensitive to variations in training and testing data. In contrast, the CNN trained on CQT features exhibited much lower and less stable results, with a mean accuracy of only 58.04%, mean precision of 54.28%, mean recall of 58.04%, and mean F1-score of 53.28%. Furthermore, the CQT-based model showed extremely high variability, with accuracy values fluctuating between 11.29% and 94.32% (standard deviation 40.71%), suggesting a lack of robustness and reliability. This instability suggests that the CQT-based model cannot be considered consistently reliable.

Table 2. Performance Comparison of CNN Model Using MFCC and CQT Features (15-Second Duration)

Scenario	Mean Accuracy	Mean Precision	Mean Recall	Mean F1 Score
Feature CQT - Duration 15s	0.580354391	0.542802606	0.580354391	0.532779325
Feature MFCC - Duration 15s	0.946719312	0.949684948	0.946719312	0.946315422

Table 3. CNN Model Fold-wise Evaluation Using MFCC and CQT Features (15s Duration)

Scenario	Fold	Accuracy	Precision	Recall	F1-Score
Feature MFCC - Duration 15s	1	0.943502825	0.945945295	0.943502825	0.943012498
Feature MFCC - Duration 15s	2	0.943502825	0.944361846	0.943502825	0.943235564
Feature MFCC - Duration 15s	3	0.965909091	0.966765343	0.965909091	0.965685843
Feature MFCC - Duration 15s	4	0.965909091	0.968859649	0.965909091	0.965451115
Feature MFCC - Duration 15s	5	0.914772727	0.922492608	0.914772727	0.914192089
Feature CQT - Duration 15s	1	0.175141243	0.030674455	0.175141243	0.052205563
Feature CQT - Duration 15s	2	0.11299435	0.012767723	0.11299435	0.022943015
Feature CQT - Duration 15s	3	0.738636364	0.784914464	0.738636364	0.715671214
Feature CQT - Duration 15s	4	0.943181818	0.948362835	0.943181818	0.941789477
Feature CQT - Duration 15s	5	0.931818182	0.937293553	0.931818182	0.931287357

Significant findings

A key finding is that MFCC is clearly more suitable than CQT for classifying Selonding gamelan compositions under the current setup. MFCC features produced both high accuracy (94.67%) and very low variance between folds (2.11% standard deviation), while CQT features led to much lower average accuracy (58.04%) and highly unstable behavior across folds. This indicates that MFCC captures discriminative timbral information that is important for distinguishing Selonding pieces, whereas the CQT representation in this configuration does not provide a robust basis for reliable classification.

Discussion

Interpretation and evaluation of the results

The results indicate that an MFCC-based CNN is already reaching practically useful performance for identifying Selonding compositions, even when the input is limited to short 15-second audio segments. The stability of the MFCC model in the range of approximately 91–97% accuracy across all five folds shows that the model is learning consistent acoustic patterns rather than memorizing individual recordings. By contrast, the dramatic fluctuation in the CQT-based model (from 11.30% to 94.32% accuracy depending on the fold) suggests that not all spectral representations are equally compatible with this CNN architecture, especially under conditions of limited training data.

Relationship between the research results and previous research or theory

The results of this study are consistent with previous research demonstrating the effectiveness of Convolutional Neural Networks (CNN) in audio classification using spectral features such as Mel-Frequency Cepstral Coefficients (MFCC). Several prior studies have shown that the combination of CNN and MFCC can effectively capture complex acoustic patterns and improve classification accuracy through techniques such as data augmentation [32],[33],[34]. Other studies have also indicated that CNNs excel in learning intricate musical patterns and enhancing sound recognition performance [35],[36]. In line with these findings, the present study confirms that a CNN trained with MFCC features can accurately classify Gamelan Selonding recordings, thereby extending the application of deep learning to non-Western traditional music and contributing to the preservation and analysis of Balinese musical heritage.

Practical implications of the research results

Practically, the MFCC-based CNN model can be used to support digital archiving and cataloging of Selonding gamelan recordings. Field recordings from ceremonies or rehearsals could be segmented into 15-second clips and automatically labeled with the most likely composition name, with typical accuracy around 94–96%. This can accelerate metadata generation in cultural archives, assist ethnomusicologists in locating specific compositions within long recordings, and help cultural institutions document and preserve repertoire without requiring expert human annotation at every initial step. Finally, the findings of this study contribute to broader technological innovation by demonstrating that deep learning architectures can be adapted to non-Western musical structures. This not only enriches research in the field of music information retrieval but also promotes the use of artificial intelligence for cultural sustainability, ensuring that sacred and ancient musical traditions such as Gamelan Selonding remain recognized, accessible, and preserved for future generations.

Limitations of the research

This study has several limitations that affect how the results should be interpreted. First, the classification task is restricted to 10 known Selonding compositions, so the generalization of the model to additional or rarer pieces has not yet been tested. Second, all evaluation was performed on fixed 15-second segments rather than full-length performances, so the behavior of the model on full recordings is still unknown. Third, the recordings come from a limited set of performance and recording conditions; differences in playing style, ensemble membership, or recording environment (e.g., village context versus studio-like capture) may alter the acoustic profile. Fourth, the approach relies entirely on supervised learning and does not yet incorporate transfer learning or few-shot learning strategies, which are often necessary for rare traditional music with very limited data.

Recommendations for further research

Future work should evaluate whether more advanced architectures, such as Convolutional Recurrent Neural Networks (CRNNs) or Transformer-based audio models, can either match or exceed the ~94–96% accuracy of the MFCC-based CNN while maintaining fold-to-fold stability. It is also important to test generalization using field recordings captured in different acoustic environments, to confirm that high performance on the current dataset does not collapse under noisy real-world conditions. In addition, future experiments should explore few-shot learning and transfer learning so that the system can recognize new or less-documented Selonding compositions from only a handful of labeled examples, which is a realistic scenario for endangered musical repertoires.

4. Conclusion:

Summary of the results and discussions

A CNN trained on MFCC features classified 10 Selonding gamelan compositions with a mean accuracy of 94.67% and a mean F1-score of 94.63% (≈2% standard deviation), while the CQT-based model reached only 58.04% mean accuracy with very high variability (40.71%). This shows that MFCC is a more reliable representation than CQT, and that even 15-second audio segments contain enough information for robust classification.

Answers to research questions or hypotheses

Yes, Selonding gamelan compositions can be automatically classified from short (15-second) audio excerpts using a CNN, as long as MFCC features are used. The results confirm that feature representation strongly influences performance, with MFCC yielding high and stable accuracy while CQT is less reliable.

Research contributions

This work provides a quantitative baseline (~94–96% per fold using MFCC) for automatic classification of Selonding gamelan, demonstrates that short segments are sufficient for identification, and shows how deep learning can support digital preservation of Balinese traditional music.

Recommendations for further research or practice

Future work should include more compositions, more recording conditions, and few-shot or transfer learning to handle rare pieces with limited data. Practically, the MFCC-based CNN can already be integrated into digital archiving workflows to assist curators and educators in labeling Selonding recordings.

Acknowledgments:

The author would like to express sincere gratitude to all individuals and institutions who have supported the completion of this research. Special appreciation is extended to the academic advisors for their invaluable guidance and constructive feedback throughout the research process. The author also thanks fellow researchers and colleagues who provided technical assistance and insights related to the implementation of the Convolutional Neural Network (CNN) model. Finally, heartfelt thanks are given to the cultural practitioners and community members who contributed to the understanding and preservation of Gamelan Selonding music, which served as the foundation of this study.

References:

[1] Y. R. Sidjabat dan J. D. Krishnanandayani, "the Transformation of Balinese Gamelan: Authenticity and Heritage Politics in Digital Platforms," *Proceeding Bali-Bhuwana Waskita Glob. Art Creat. Conf.*, vol. 4, hal. 408–414, 2024, doi: 10.31091/bbwp.v4i1.610

[2] K. D. Pradnyani, "Sistem Klasifikasi Gamelan Bali Berbasis Web a Website-Based Balinese Classification System of the Balinese gamelan," vol. 16, no. 2, hal. 82–90, 2023

[3] S. P. Collins *et al.*, "Sonic Traditions: Exploring The Gamelan's Influence On Contemporary Indonesian Music," vol. 10, no. 19, hal. 167–186, 2021

[4] I. K. Ardana, "Re-Actualization Balinese Gamelan Harmony for Renewal Knowlegde of the Balinese Music," *Int. J. Creat. Arts Stud.*, vol. 8, no. 1, hal. 51–69, 2021, doi: 10.24821/ijcas.v8i1.5514

[5] K. A. W. Pradana, I. W. Rai S, dan I. W. Suherta, "The The Musicality of Gamelan Gong Kebyar Mepacek as a North Bali Traditional Music Identity," *Randwick Int. Soc. Sci. J.*, vol. 4, no. 2, hal. 241–253, 2023, doi: 10.47175/rissj.v4i2.653

[6] I. W. Rai, N. M. Ruastiti, G. Y. K. Pradana, dan Y. Wafom, "Gamelan Selonding As Part of an Essential Instrument in the Sustainment of the Sacred Art Activities of the Batur Indigenous People in Kintamani, Bali," *J. Law Sustain. Dev.*, vol. 11, no. 5, hal. 1–22, 2023, doi: 10.55908/sdgs.v11i5.1006

[7] V. Wulansari dan S. Maisy, "Kelestarian Budaya Dan Adat Di Desa Tengan Pegringsingan Karangasem Bali," *J. Fash.*, vol. 1, no. 2, hal. 5, 2023.

[8] K. A. T. Paramitha, I. W. D. Putra, dan N. W. Mudiasih, "Pengembangan Aplikasi Android Pembelajaran Gamelan Selonding Gaya Tenganan," *J. Music Sci. Technol. Ind.*, vol. 5, no. 2, hal. 223–239, 2022, doi: 10.31091/jomsti.v5i2.2134

[9] A. Agung, G. Agung, dan R. Putra, "Kemampuan Menabuh Gending Rejang Manda Dalam Gamelan Selonding Gaya Bebandem Oleh Komunitas Selonding Bali Aga Banjar Pande Tunggak Bebandem Karangasem Tahun 2022," vol. II, no. April, 2022, doi: 10.5281/zenodo.7302722

[10] I. W. Y. M. Putra, "Development Of Sekaa Selonding Manik Selukat Banjar Tunjuk Kelod | Pembinaan Sekaa Selonding Manik Selukat Banjar Tunjuk Kelod," *GHURNITA J. Seni Karawitan*, vol. 3, no. 4, hal. 372–378, 2023, doi: 10.59997/jurnalsenikarawitan.v3i4.2495

[11] I. W. Sudiarsa, "Gamelan Salonding Masuk Hotel Studi Kasus: Perkembangan Gamelan Salonding di Desa Tenganan Pegringsingan, Kecamatan Manggis …," *Widyanatya*, vol. 3, hal. 31–36, 2021.

[12] I. G. Harsemadi, "Perbandingan Kinerja Algoritma K-NN dan SVM dalam Sistem Klasifikasi Genre Musik Gamelan Bali," *INFORMATICS Educ. Prof. J. Informatics*, vol. 8, no. 1, hal. 1, 2023, doi: 10.51211/itbi.v8i1.2417

[13] Firman, Firdaus, M. Halim, Alfalah, dan Sriyanto, "Analisis Pola Musik Karawitan di Tengah Era Digital," *Indones. J. Comput. Sci.*, vol. 13, no. 2, hal. 3157–3167, 2024, doi: 10.33022/ijcs.v13i2.3783

[14] W. Vitale dan W. Sethares, "Balinese Gamelan Tuning: The Toth Archives," *Anal. Approaches to World Music*, vol. 9, no. 2, 2021

[15] V. L. Hardjanto dan Wahyono, "Audio Enhancement for Gamelan Instrument Recognition using Spectral Subtraction," *Eng. Technol. Appl. Sci. Res.*, vol. 15, no. 2, hal. 22042–22048, 2025, doi: 10.48084/etasr.10181

[16] Y. Mardiati, . W., dan D. R. Puguh, “Making Connection: Integrating Gamelan Music into Social Studies Classroom,” *Int. J. Res. Rev.*, vol. 8, no. 7, hal. 236–244, 2021, doi: 10.52403/ijrr.20210733

[17] Z. Liu, “Audio Feature Extraction and Classification Technology Based on Convolutional Neural Network,” *J. Electr. Syst.*, hal. 1425–1431, 2024

[18] R. Jahangir, M. A. Nauman, R. Alroobaea, J. Almotiri, M. M. Malik, dan S. M. Alzahrani, “Deep Learning-based Environmental Sound Classification Using Feature Fusion and Data Enhancement,” *Comput. Mater. Contin.*, vol. 74, no. 1, hal. 1069–1091, 2023, doi: 10.32604/cmc.2023.032719

[19] I. G. Agung, C. Wijaya, I. G. A. Indrawan, I. N. A. Fajaraditya, dan I. Gusti, “Classification Of Bougainvillea Flower Varieties Using Variant Of CNN : Resnet50,” vol. 6, no. 2, hal. 154–162, 2025

[20] M. Oktaviani, T. E. Sutanto, dan M. Mahmudi, “Klasifikasi Usia Berdasarkan Suara Dengan Ekstraksi Ciri Mel Frequency Cepstral Coefficients Menggunakan Support Vector Machine,” *Kesatria J. Penerapan Sist. Inf. (Komputer dan Manajemen)*, vol. 4, no. 4, hal. 901–907, 2023.

[21] G. Abou Haidar, “Music Genre Classification Using Adam Algorithm of Convolutional Neural Network,” *Int. J. Inf. Commun. Technol.*, vol. 10, no. 2, hal. 152–160, 2024, doi: 10.21108/ijoict.v10i2.978

[22] R. Luis dan N. Rokhman, “Traditional Music Regional Classification using Convolutional Neural Network (CNN),” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 16, no. 4, hal. 379, 2022, doi: 10.22146/ijccs.73910

[23] R. Soerkarta, S. Aras, dan A. N. Aswad, “Hyperparameter Optimization of CNN Classifier for Music Genre Classification,” *J. RESTI*, vol. 7, no. 5, hal. 1205–1210, 2023, doi: 10.29207/resti.v7i5.5319

[24] Royan Hisyam Rafliansyah, Basuki Rahmat, dan Chrystia Aji Putra, “Klasifikasi Suara Instrumen Musik Tiup Menggunakan Metode Convolutional Neural Network,” *Merkurius J. Ris. Sist. Inf. dan Tek. Inform.*, vol. 2, no. 4, hal. 01–09, 2024, doi: 10.61132/merkurius.v2i4.119

[25] S. Y. Yusdiantoro dan T. B. Sasongko, “Implementasi Algoritma MFCC dan CNN dalam Klasifikasi Makna Tangisan Bayi,” *Indones. J. Comput. Sci.*, vol. 12, no. 4, hal. 1957–1968, 2023, doi: 10.33022/ijcs.v12i4.3243

[26] S. Ihsanti *et al.*, “Klasifikasi Genus Burung Hantu Berdasarkan Suara Menggunakan Convolutional Neural Network,” vol. 11, no. 4, hal. 4532–4538, 2024

[27] R. F. Fadhillah dan R. Sumiharto, “Klasifikasi Suara Untuk Memonitori Hutan Berbasis Convolutional Neural Network,” *IJEIS (Indonesian J. Electron. Instrum. Syst.*, vol. 13, no. 1, 2023, doi: 10.22146/ijeis.79536

[28] F. Ferdiawan dan B. Hartono, “Deteksi Suara Chord Piano Menggunakan Metode,” vol. 5, no. 1, hal. 62–68, 2022

[29] F. Wolf-Monheim, “Spectral and Rhythm Features for Audio Classification with Deep Convolutional Neural Networks,” 2024, [Daring]. Tersedia pada: http://arxiv.org/abs/2410.06927

[30] Z. Meng dan W. Chen, “Automatic music transcription based on convolutional neural network, constant Q transform and MFCC,” *J. Phys. Conf. Ser.*, vol. 1651, no. 1, 2020, doi: 10.1088/1742-6596/1651/1/012192

[31] J. Yang *et al.*, “A Deep-Learning Framework with Multi-Feature Fusion and Attention Mechanism for Classification of Chinese Traditional Instruments,” *Electron.*, vol. 14, no. 14, hal. 1–18, 2025, doi: 10.3390/electronics14142805

[32] H. Train, “applied sciences A Binaural MFCC-CNN Sound Quality Model of,” 2022.

[33] Pratibha Rashmi dan Manu Pratap Singh, “Convolution neural networks with hybrid feature extraction methods for classification of voice sound signals,” *World J. Adv. Eng. Technol. Sci.*, vol. 8, no. 2, hal. 110–125, 2023, doi: 10.30574/wjaets.2023.8.2.0083

[34] H. C. Chu, Y. L. Zhang, dan H. C. Chiang, “A CNN Sound Classification Mechanism Using Data

Augmentation," *Sensors*, vol. 23, no. 15, 2023, doi: 10.3390/s23156972

[35] N. Asanah dan I. Pratama, "Deep Learning Approach for Music Genre Classification using Multi-Feature Audio Representations," *Sistemasi*, vol. 14, no. 5, hal. 2045, 2025, doi: 10.32520/stmsi.v14i5.5369

[36] A. Wirdiani, S. N. Machetho, I. K. G. D. Putra, M. Sudarma, R. S. Hartati, dan H. A. Ferdian, "Improvement Model for Speaker Recognition using MFCC-CNN and Online Triplet Mining," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 14, no. 2, hal. 420–427, 2024, doi: 10.18517/ijaseit.14.2.19396