Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14019
e-ISSN : 2541-2019
p-ISSN : 2541-044X

# Performance Comparison of KNN and CNN in Classifying Balinese *Gangsa* Instrument Tones

**I Gede Putra Mas Yusadara[1)*], Ni Made Rai Masita Dewi[2)], I Gede Bintang Arya Budaya[3)]**
[1,2,3)] Institute of Technology and Business STIKOM Bali, Indonesia
[1)]putramas@stikom-bali.ac.id , [2)]raimasita@stikom-bali.ac.id , [3)]bintang@stikom-bali.ac.id

**Submitted** :Aug 14, 2024 | **Accepted** : Sept 19, 2024 | **Published** : Oct 3, 2024

**Abstract:** Balinese traditional music, particularly the Gamelan Gangsa, represents a unique aspect of Indonesia's cultural heritage. Despite its cultural significance, the study and teaching of this instrument face challenges, particularly in tone standardization and the availability of effective learning tools. This research addresses these challenges by exploring the application of Artificial Intelligence (AI) technologies specifically K-Nearest Neighbors (KNN) and Convolutional Neural Networks (CNN) in the identification and classification of Gamelan Gangsa tones. The study involved the creation of a dataset comprising audio recordings of the instrument, followed by the development and evaluation of KNN and CNN models. The results indicate that KNN, with an accuracy of 90%, outperformed CNN, which achieved an accuracy of 85%. The findings suggest that KNN is particularly effective in distinguishing subtle tonal differences, making it a valuable tool for supporting traditional music education. This research not only contributes to the technical understanding of Gamelan Gangsa's acoustic characteristics but also underscores the potential of AI in cultural preservation. The development of AI-based tone identification systems can facilitate the teaching and learning of traditional music, ensuring its transmission to future generations. The study serves as a foundation for further exploration into the integration of AI technologies with cultural heritage, demonstrating how modern innovations can enhance the appreciation and understanding of traditional arts.

**Keywords:** Audio Signal Processing; Traditional Music Instrument; Tones Identification; Chroma Features; MFCCs.

## INTRODUCTION

Balinese traditional music, particularly that produced by the Gamelan Gangsa instrument, is a significant cultural heritage, reflecting the richness and uniqueness of Indonesia's musical traditions (McGraw, 2008; Pradana & Suherta, 2023; Rosmiati et al., 2024). The practice of learning Gamelan instruments is commonly conducted by art studios in Bali, with almost every community in a village possessing a complete set of Balinese Gamelan instruments. The first challenge in this learning process is that instructors cannot facilitate all participants simultaneously. The second challenge is the standardization of the instrument's tones, as expert opinions can vary, even in the smallest tonal discrepancies. Consequently, there is a need for methods that can accurately and efficiently identify and evaluate the tones of these instruments.

In modern musical instruments such as the guitar and piano, numerous studies have already developed smart learning tools (Boestad & Rudberg, 2021; Kulkarni et al., 2020; Kumar et al., 2021; Li, 2022; Liu & Huang, 2021). However, for traditional instruments like the Gamelan Gangsa, research in this area is still minimal. The lack of innovative and adaptive learning resources hinders the introduction of this instrument's uniqueness to a broader audience and the younger generation (Mukti & Fathurrahman, 2023).

Artificial Intelligence (AI) technologies, such as K-Nearest Neighbors (KNN) and Convolutional Neural Networks (CNN), have proven effective in various applications, including sound processing and pattern recognition (Demir et al., 2020; Prabavathy et al., 2020; Tsalera et al., 2020). The implementation of these technologies in music studies, particularly in the identification of traditional music instrument tones, offers revolutionary potential. By leveraging KNN and CNN, this research aims to develop a system capable of recognizing and classifying Gamelan Gangsa tones, thereby supporting preservation efforts and facilitating the teaching and learning of traditional Balinese music.

While there have been studies related to the application of AI in music, few have focused on applying this technology to traditional musical instruments, specifically the Balinese Gamelan Gangsa. Interpreting the complexities of traditional music tones with high accuracy remains a challenge. Therefore, further exploration is

*I Gede Putra Mas Yusadara

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14019
e-ISSN : 2541-2019
p-ISSN : 2541-044X

needed to determine how KNN and CNN can be adapted and optimized to address the unique challenges faced in traditional music studies.

This research aims to design an identification model capable of accurately identifying Gamelan Gangsa instrument tones. Through the application of KNN and CNN, this study seeks to enhance technical understanding of the acoustic characteristics of Gamelan Gangsa, develop a tone identification system that can serve as a learning tool, and contribute to the preservation and promotion of Balinese traditional music through modern technology. The results of this research are expected to be a starting point in the identification of traditional music tones, which can be used as a benchmark for future technological developments. This is an important step in integrating technology and cultural preservation, demonstrating how technological innovation can be leveraged to support cultural heritage (Ahmad et al., 2021; Parker & Spennemann, 2022; Sudipa et al., 2022).

## LITERATURE REVIEW

The application of artificial intelligence (AI) and machine learning in sound classification has seen significant advancements, offering valuable insights for various domains, including music education and health diagnostics. These advancements provide a strong foundation for exploring the identification of tones in traditional musical instruments, such as the Balinese Gamelan Gangsa.

One study explored the integration of deep learning (DL) and AI in enhancing music education for children, particularly in the context of piano instruction. By utilizing Convolutional Neural Networks (CNN) to detect the onset of piano notes, the research demonstrated a significant improvement in learning outcomes. The implementation of this method in teaching preschoolers not only enhanced their engagement with the music but also received positive feedback from parents, who noted a marked increase in their children's interest in learning (Li, 2022). This study highlights the potential of DL and AI to create more interactive and engaging learning experiences in music education.

Another study focused on the development of a piano teaching system that employs neural network technology to make instruction more accessible and interactive. This system addresses the limitations of traditional teaching methods, which often require direct, face-to-face interaction between teacher and student. By evaluating the students' piano playing and providing real-time guidance, the neural network-based system offers a new approach to music education that is both cost-effective and engaging (Liu & Huang, 2021). This work underscores the versatility of AI in transforming conventional educational practices.

In the field of health diagnostics, a study on heart sound classification combined Mel-Frequency Cepstrum Coefficient (MFCC) feature extraction with a Multi-Scale Residual Recurrent Neural Network (MsRes-RNN) model. This approach was applied to classify heart sounds, achieving a high classification accuracy of 93.9% using the PhysioNet/CinC Challenge 2016 dataset (Chen et al., 2022). The success of this method in accurately classifying complex sound patterns demonstrates the effectiveness of combining advanced feature extraction techniques with sophisticated neural network models.

Additionally, research in audio classification has explored methods to differentiate between speech and music by leveraging both visual and spectral features of chromagrams. This study utilized eigenvector-based centrality for feature selection, which significantly enhanced the classification accuracy when tested on various audio databases. The findings indicated that chromagram texture descriptors were particularly effective in distinguishing between different types of audios, even under challenging conditions where training and testing datasets did not match (Birajdar & Patil, 2020). This study illustrates the robust potential of chromagram-based features in audio classification tasks.

While these studies demonstrate the broad applicability of AI and machine learning in sound classification, there remains a notable gap in their application to traditional music instruments, such as the Balinese Gamelan Gangsa. The unique acoustic properties and cultural significance of such instruments require tailored approaches that are not yet widely explored in the existing literature.

This research aims to address this gap by developing an identification model that can accurately recognize and classify the tones of the Gamelan Gangsa. By adapting the methodologies and insights gained from previous studies, this research seeks to contribute to the preservation and promotion of Balinese traditional music, while also exploring the potential of AI in bridging the gap between technology and cultural heritage.

## METHOD

The method used in this research is shown in the Figure 1 below.

*I Gede Putra Mas Yusadara

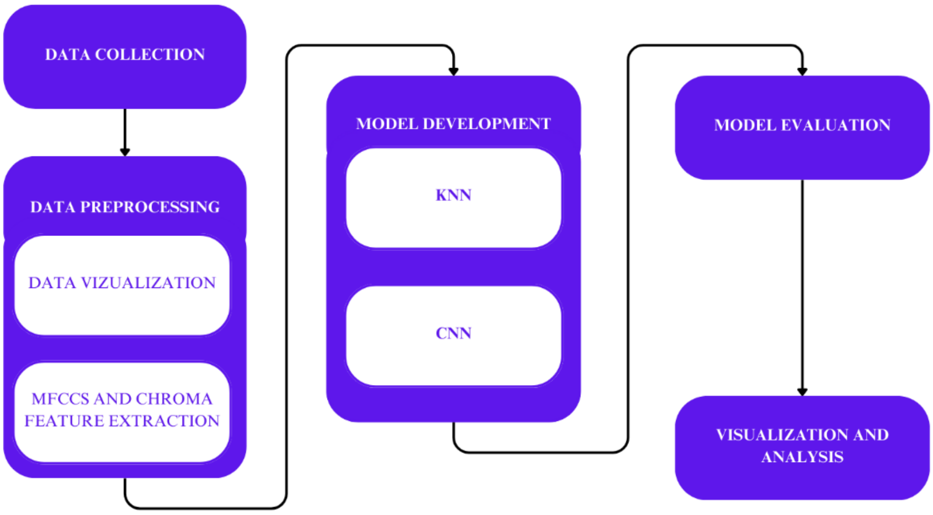


Fig.1 Research method

***Data Collection***

The study begins with the systematic collection of a dataset consisting of audio recordings of the Gamelan Gangsa, a traditional Balinese musical instrument. Each recording is meticulously labeled to indicate the specific pitch it represents. The dataset includes metadata such as recording conditions and the equipment used, which are critical for contextualizing the data. The final collected data contains 50 recordings of each class, with a total of 500 recordings. Figure 2 shows the dataset collection process and Figure 3 shows the saved dataset in Google Drive.

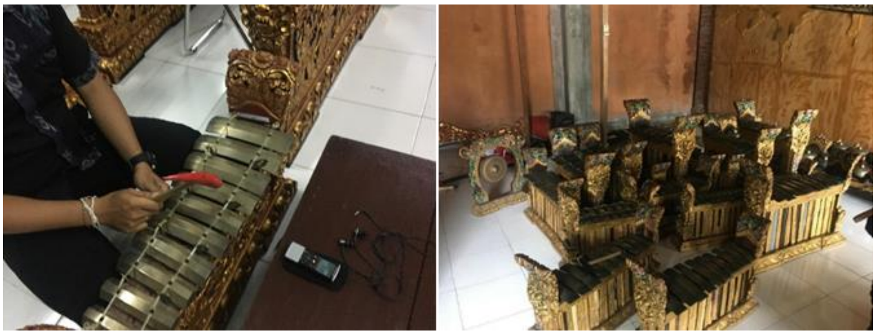

Fig.2 The process of collecting the dataset of *Gangsa* tones

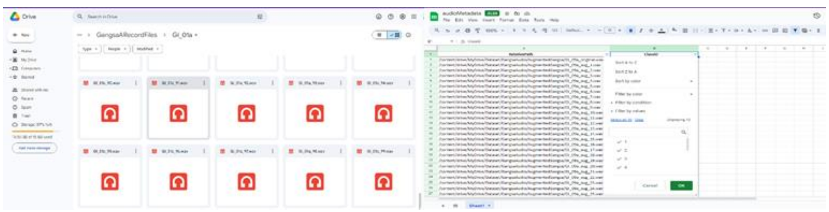

Fig.3 Saved *Gangsa* tone recordings and metadata in Google Drive

***Data Preprocessing***

During preprocessing, the audio data is processed using advanced feature extraction techniques, including Mel-Frequency Cepstral Coefficients (MFCC) (Abdul & Al-Talabani, 2022; Birajdar & Patil, 2020; Tiwari et al., 2020), and Chroma features (Birajdar & Patil, 2020; Su et al., 2020), which are essential for capturing the unique acoustic characteristics of the *Gangsa*. To ensure consistency, the extracted features are normalized and padded to uniform sizes. This step is crucial for maintaining data integrity during model training. The dataset is then divided into training and testing sets using stratified sampling to preserve class distribution, ensuring each pitch class is adequately represented across both sets. Figure 4 shows a sample sound waveform of Gangsa tones, Figure 5 shows the sample visualization of MFCCs, and Figure 6 shows the sample visualization of the Chroma feature.

*I Gede Putra Mas Yusadara

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14019
e-ISSN : 2541-2019
p-ISSN : 2541-044X

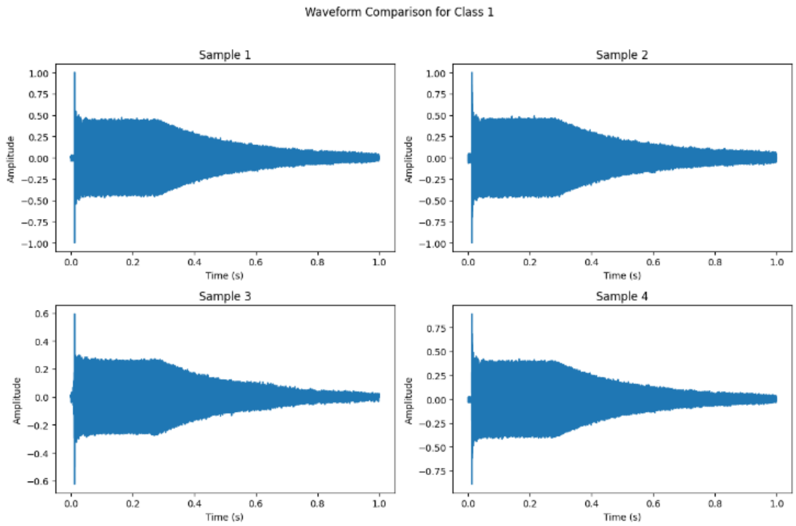


Fig.4 Sample visualization of sound waveforms from 4 Gangsa tone recordings in Class 1

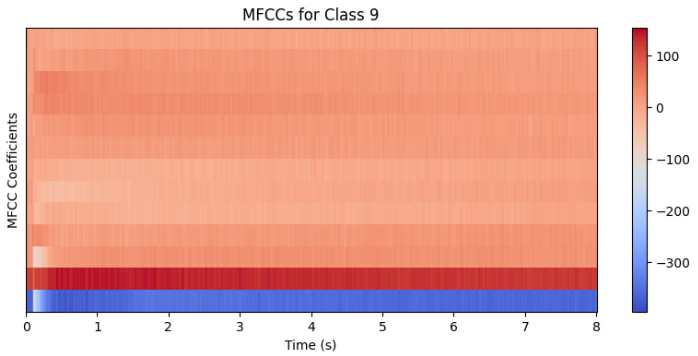


Fig.5 Sample visualization of MFCCs from one of the Class 9 tones recordings

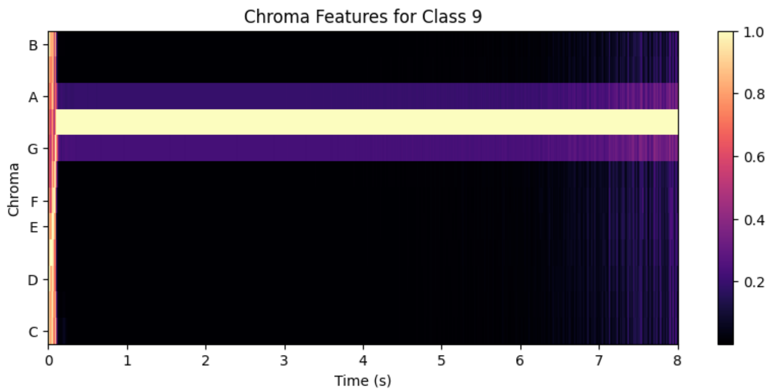


Fig.6 Sample visualization of the Chroma feature from one of the Class 9 tone recordings

***Model Development***

Two distinct models are developed for pitch identification, CNN (Das et al., 2020) and KNN model. The CNN architecture is specifically designed with layers of convolutional and pooling operations, followed by fully connected layers. The model learns hierarchical features from the audio data, such as local patterns in the sound waves, which are then flattened to serve as input for the fully connected layers. The CNN model is trained on the stratified training set, with loss and accuracy metrics tracked over several epochs to monitor learning progression. In parallel, the KNN model is developed, where the features are flattened to be compatible with the algorithm. The KNN model is trained on the same dataset, facilitating a direct performance comparison with the CNN model.

***Model Evaluation***

The CNN and KNN models are evaluated using comprehensive performance metrics, including accuracy, precision, recall, and F1-score (Powers, 2020; Yacouby & Axman, 2020). These metrics provide a detailed assessment of each model's ability to classify audio samples correctly. Confusion matrices are generated for both

*I Gede Putra Mas Yusadara

e-ISSN : 2541-2019
p-ISSN : 2541-044X

models, offering insights into their performance on a class-by-class basis. These matrices are visualized to highlight areas of strength and weakness, enabling a nuanced understanding of model performance. A comparative analysis is conducted, using bar plots to visually represent the differences in the models' performance metrics. This comparison helps identify which model is better suited for the task of pitch identification in the Gamelan Gangsa and informs decisions about potential improvements.

$$\text{Accuracy} = \frac{\text{True Positives+True Negatives}}{\text{Total Observations}} \quad (1)$$

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives+False Positives}} \quad (2)$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives+False Negatives}} \quad (3)$$

$$\text{F1 Score} = 2 * \frac{\text{Precision}*\text{Recall}}{\text{Precision+Recall}} \quad (4)$$

***Visualization and Analysis***

To enhance the interpretability of the findings, various visualizations are created throughout the analysis. The class distribution within the dataset is plotted to provide an overview of the data composition and to confirm that the stratified split maintains class balance in the training and testing sets. The training progress of the CNN model is visualized by plotting loss and accuracy metrics over the epochs, allowing for an assessment of the model's learning dynamics. Further visualizations include confusion matrices and classification reports for both the CNN and KNN models, which are crucial for understanding the detailed performance across different classes. Finally, a comparative analysis is conducted using bar plots to juxtapose the key performance metrics of the CNN and KNN models, providing a clear summary of their relative strengths and areas for improvement.

## RESULT

The KNN model implementation begins by flattening the feature data to make it compatible with the KNN algorithm, which requires a 2D input format. The features are then normalized using *Standard Scaler* to ensure that all feature values are on the same scale, which is crucial for distance-based algorithms like KNN. The KNN model is initialized with 5 neighbors, using the Manhattan distance metric and a distance-based weighting scheme, where closer neighbors have a greater influence on the prediction. After training the model on the flattened and normalized data, predictions are made on the test set. The model's performance is evaluated using accuracy, precision, recall, and F1 score, providing a comprehensive assessment of how well the KNN model classifies the audio samples. The results are then printed, highlighting the effectiveness of the model with the chosen settings.

The CNN model implementation begins by designing a custom architecture that includes multiple convolutional layers followed by pooling layers and fully connected layers. The input features, which are derived from MFCC, and Chroma features, are passed through these layers to automatically learn hierarchical patterns in the data. The convolutional layers capture local features from the audio signals, while the pooling layers help in reducing the dimensionality. The output from the final convolutional layer is flattened to create a feature vector, which is then fed into the fully connected layers. The model is trained using the training dataset, with the loss and accuracy monitored over several epochs to ensure the model is learning effectively. During training, dropout is applied to prevent overfitting by randomly setting a fraction of input units to zero during each update cycle. After training, the model's performance is evaluated on the test set using metrics such as accuracy, precision, recall, and F1 score. These metrics provide a detailed understanding of the model's ability to accurately classify the different audio classes. The final results are used to assess the overall effectiveness of the CNN model in handling the pitch recognition task for the Gamelan *Gangsa* instrument. During training, the model implemented early stopping, which resulted in the training ending after 8 epochs. Figure 7 shows the training loss and accuracy for the CNN model.

*I Gede Putra Mas Yusadara

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14019
e-ISSN : 2541-2019
p-ISSN : 2541-044X

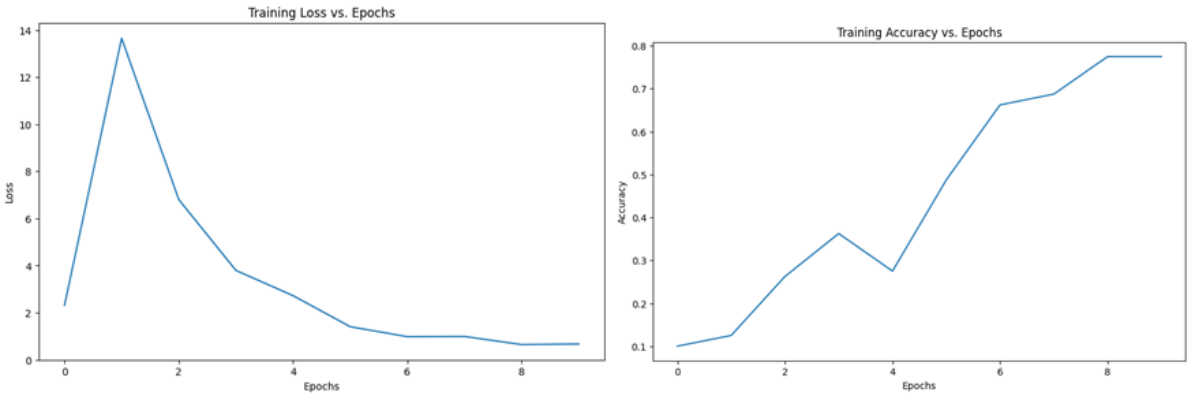


Fig.7 Training loss and accuracy during epochs CNN model

As shown in Table 1 and Figure 8, the results from the KNN and CNN models evaluation, reveal distinct differences in performance. The KNN model achieved an accuracy of 90%, indicating that it was able to correctly classify the majority of the audio samples. Its precision, recall, and F1 score are also notably high, at 85%, 90%, and 86.67% respectively. This suggests that the KNN model is effective at identifying the correct classes, with a relatively low rate of false positives and false negatives. The confusion matrix for KNN shows that most predictions are concentrated on the diagonal, which is expected in a well-performing model, indicating that the majority of the predictions match the true labels. In contrast, the CNN model, while also performing well, exhibits slightly lower accuracy at 85%. The precision and recall for CNN are higher than the accuracy, at 88.33% and 85% respectively, with an F1 score of 84.33%. This balance between precision and recall indicates that the CNN model is more cautious in its predictions, potentially reducing false positives but at the cost of missing some true positives. The CNN confusion matrix shows more variation off the diagonal, particularly with a few classes where the model confuses between similar classes. This suggests that while the CNN model is generally reliable, it may struggle with certain classes that have subtle differences in their audio features.

Table 1. KNN and CNN evaluation comparison

| No | Model | Accuracy | Precision | Recall | F1-Scores |
|---|---|---|---|---|---|
| 1 | KNN | **0,90** | 0,85 | **0,90** | **0,87** |
| 2 | CNN | 0,85 | **0,88** | 0,85 | 0,84 |

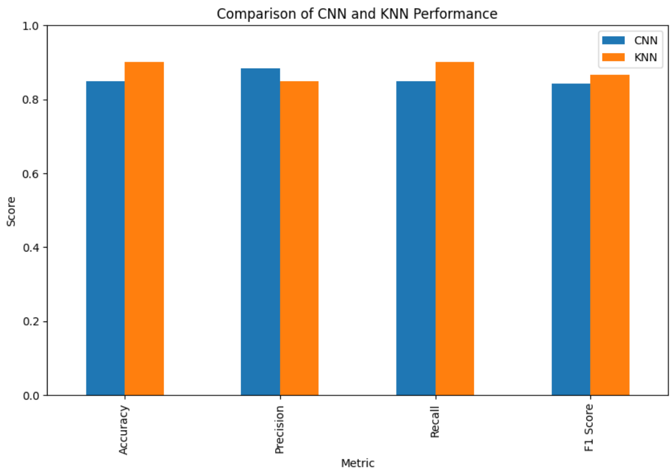


Fig.8 CNN and KNN comparison

Overall, the KNN model appears to outperform the CNN model in this particular task, especially in terms of accuracy and recall. However, the CNN model's slightly higher precision indicates it may be better at avoiding incorrect classifications, which could be valuable in certain applications. These results highlight the importance of choosing the right model based on the specific needs of the task, as each model offers different strengths. The KNN model's performance could be attributed to the effective feature normalization and possibly the use of distance-based weighting, which makes it well-suited for this type of classification task. Meanwhile, the CNN

*I Gede Putra Mas Yusadara

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Number 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14019

e-ISSN : 2541-2019
p-ISSN : 2541-044X

model, despite its slight performance lag, benefits from its ability to learn complex feature hierarchies, which could be advantageous with further tuning or with more data.

## DISCUSSIONS

The analysis of the confusion matrices reveals that the CNN model, while generally effective, struggles with certain classes, leading to some misclassification. Specifically, classes such as Class 2 and Class 5 show instances where the model confuses them with other similar classes. This misclassification could be attributed to the nature of the audio features extracted from these classes. In traditional Balinese music, slight variations in pitch and timbre between notes can be difficult for a model to distinguish, especially if the training data does not fully capture these subtle differences. The CNN, as shown in Figure 9, despite its ability to learn complex patterns, might not be able to fully differentiate these classes due to the similar spectral features in the MFCCs, and Chroma features, leading to these errors.

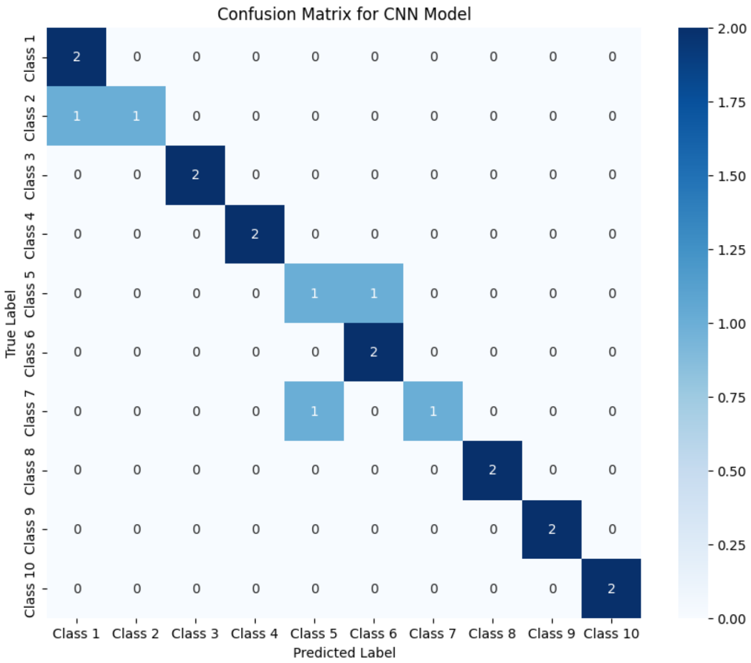


Fig.9 Confusion matrix of CNN Model

On the other hand, the KNN model, as shown in Figure 10, demonstrates a higher accuracy and recall, which suggests it is better at distinguishing between the different classes of Gamelan Gangsa. One possible reason for KNN's superior performance in this case is its reliance on distance metrics, which can be highly effective when the feature space is well-normalized and the features are well-separated. The use of the Manhattan distance metric, combined with distance-based weighting, allows KNN to give more importance to closer neighbors, which may be crucial in distinguishing similar audio features. This approach can outperform a neural network like CNN, particularly when the dataset is not large enough for the CNN to fully learn the complex feature hierarchies or when the features themselves are well-represented in a lower-dimensional space where KNN excels.

*I Gede Putra Mas Yusadara

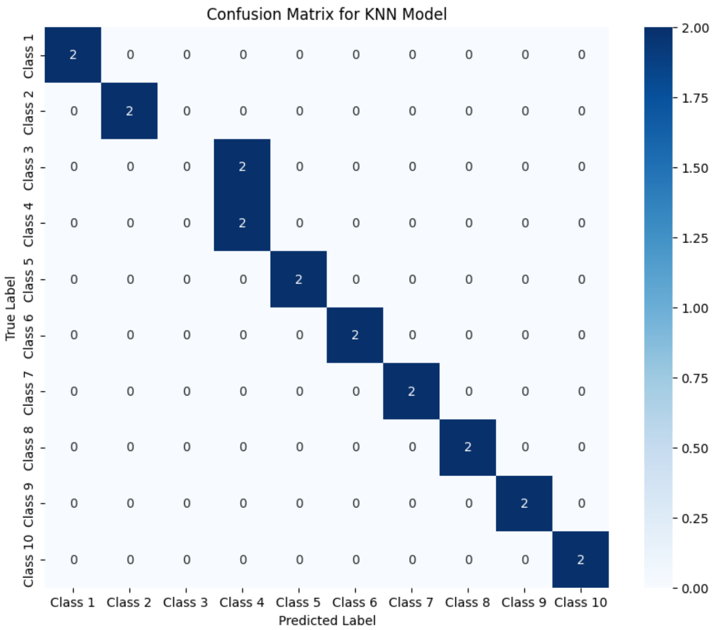


Fig.10 Confusion matrix of KNN Model

Moreover, KNN's performance might also benefit from the relatively small dataset size and the use of standardization. In smaller datasets, KNN can perform better because it does not rely on a large amount of data to learn model parameters as CNN does. Instead, KNN directly uses the training samples to make predictions, which can be advantageous when the dataset is limited and the features are properly scaled. In contrast, CNNs generally require more data to effectively learn the necessary weights and biases across their layers, and when data is limited, the model may not generalize well, leading to the observed misclassifications.

In summary, the misclassifications in the CNN model likely stem from the challenges it faces in distinguishing similar classes with subtle feature differences, exacerbated by the model's need for more extensive training data. Meanwhile, the KNN model's reliance on distance metrics and its effectiveness in a well-normalized feature space explain its superior performance in this particular task. The discussion underscores the importance of choosing the appropriate model based on the characteristics of the dataset and the specific requirements of the classification task.

## CONCLUSION

This study demonstrated the potential of applying KNN and CNN in the identification and classification of tones produced by the Balinese Gamelan Gangsa, a traditional musical instrument of significant cultural importance. By leveraging these AI technologies, the research not only highlighted the acoustic complexities inherent in traditional music but also provided a foundation for the development of innovative learning tools that can support the preservation and teaching of Balinese music. The comparative analysis revealed that while KNN outperformed CNN in this specific application, both models offer valuable insights into the potential of AI to bridge the gap between cultural preservation and modern technology. These findings underscore the importance of continued research and development in this area, aiming to create adaptive and accessible educational resources that can engage younger generations and broaden the appreciation of Indonesia's rich musical heritage.

## ACKNOWLEDGMENT

We extend our sincere thanks to the Ministry of Education, Culture, Research, and Technology, particularly the Directorate General of Higher Education, Research, and Technology, for their financial support through the Beginner Lecturer Research (*Penelitian Dosen Pemula/PDP*) Program 2024. We also express our gratitude to the Institute of Technology and Business STIKOM Bali for their support in the implementation of this research.

## REFERENCES

Abdul, Z. K., & Al-Talabani, A. K. (2022). Mel frequency cepstral coefficient and its applications: A review. *IEEE Access*, *10*, 122136–122158.

Ahmad, I., Rahmanto, Y., Pratama, D., & Borman, R. I. (2021). Development of augmented reality application for introducing tangible cultural heritages at the lampung museum using the multimedia development life cycle. *Ilk. J. Ilm*, *13*(2), 187–194.

Birajdar, G. K., & Patil, M. D. (2020). Speech/music classification using visual and spectral chromagram features. *Journal of Ambient Intelligence and Humanized Computing*, *11*(1), 329–347.

Boestad, A., & Rudberg, F. (2021). *Piezoelectric Guitar Tuner*.

*I Gede Putra Mas Yusadara

Chen, Q., Wu, Z., Zhong, Q., & Li, Z. (2022). Heart Sound Classification Based on Mel-Frequency Cepstrum Coefficient Features and Multi-Scale Residual Recurrent Neural Networks. *Journal of Nanoelectronics and Optoelectronics*, *17*(8), 1144–1153.

Das, J. K., Ghosh, A., Pal, A. K., Dutta, S., & Chakrabarty, A. (2020). Urban sound classification using convolutional neural network and long short term memory based on multiple features. *2020 Fourth International Conference On Intelligent Computing in Data Sciences (ICDS)*, 1–9.

Demir, F., Abdullah, D. A., & Sengur, A. (2020). A new deep CNN model for environmental sound classification. *IEEE Access*, *8*, 66529–66537.

Kulkarni, U., Kaushik, S., Lobo, L., & Sonkusare, R. (2020). Comparative study of digital signal processing techniques for tuning an acoustic guitar. *2020 7th International Conference on Smart Structures and Systems (ICSSS)*, 1–5.

Kumar, B. D., Kushwaha, A., Kumar, A., & Agarwal, A. (2021). Design & Implementation of Digital Guitar Tuner Using MATLAB. *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 547–549.

Li, W. (2022). Analysis of piano performance characteristics by deep learning and artificial intelligence and its application in piano teaching. *Frontiers in Psychology*, *12*, 751406.

Liu, M., & Huang, J. (2021). Piano playing teaching system based on artificial intelligence–design and research. *Journal of Intelligent & Fuzzy Systems*, *40*(2), 3525–3533.

McGraw, A. C. (2008). Different temporalities: the time of Balinese gamelan. *Yearbook for Traditional Music*, *40*, 136–162.

Mukti, R. W. T., & Fathurrahman, M. F. (2023). PENGEMBANGAN MEDIA PEMBELAJARAN ARTSTEPS UNTUK MENINGKATKAN HASIL BELAJAR SENI MUSIK: Materi Alat Musik Tradisional Siswa Kelas IV SD. *Jurnal Pendidikan Teknologi Informasi (JUKANTI)*, *6*(2), 329–341.

Parker, M., & Spennemann, D. H. R. (2022). Classifying sound: a tool to enrich intangible heritage management. *Acoustics Australia*, *50*(1), 23–39.

Powers, D. M. W. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *ArXiv Preprint ArXiv:2010.16061*.

Prabavathy, S., Rathikarani, V., & Dhanalakshmi, P. (2020). Classification of Musical Instruments using SVM and KNN. *International Journal of Innovative Technology and Exploring Engineering*, *9*(7), 1186–1190.

Pradana, K. A. W., & Suherta, I. W. (2023). The Musicality of Gamelan Gong Kebyar Mepacek as a North Bali Traditional Music Identity. *Randwick International of Social Science Journal*, *4*(2), 241–253.

Rosmiati, T., Bahari, J. P., & Yuliyanti, A. (2024). Pola Barisan Dalam Alat Musik Gangsa Daerah Bali. *Cerdika: Jurnal Ilmiah Indonesia*, *4*(2), 102–108.

Su, Y., Zhang, K., Wang, J., Zhou, D., & Madani, K. (2020). Performance analysis of multiple aggregated acoustic features for environment sound classification. *Applied Acoustics*, *158*, 107050.

Sudipa, I. G. I., Aditama, P. W., & Yanti, C. P. (2022). Developing Augmented Reality Lontar Prasi Bali as an E-learning Material to Preserve Balinese Culture. *J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl.*, *13*(4), 169–181.

Tiwari, S., Sapra, V., & Jain, A. (2020). Heartbeat sound classification using Mel-frequency cepstral coefficients and deep convolutional neural network. In *Advances in Computational Techniques for Biomedical Image Analysis* (pp. 115–131). Elsevier.

Tsalera, E., Papadakis, A., & Samarakou, M. (2020). Monitoring, profiling and classification of urban environmental noise using sound characteristics and the KNN algorithm. *Energy Reports*, *6*, 223–230.

Yacouby, R., & Axman, D. (2020). Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models. *Proceedings of the First Workshop on Evaluation and Comparison of NLP Systems*, 79–91.

*I Gede Putra Mas Yusadara