Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15494

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

Music-Structure Segmentation in Balinese Gamelan (Tabuh Lelambatan) with SSM, Checkerboard Novelty, and HMM

Ni Nyoman Sucianta Pertiwi[1)*], Anak Agung Gde Bagus Ariana[2)], Ni Putu Suci Meinarni[3)], Ayu Gede Willdahlia[4)], Made Suci Ariantini[5)]
[1,2,3,4,5)]Informatika, Fakultas Teknologi dan Informatika, Institut Bisnis dan Teknologi Indonesia, Bali, Indonesia
[1)]suciantapertiwi03@gmail.com, [2)]gungariana@instiki.ac.id, [3)]sucimeinarni@instiki.ac.id,
[4)]willdahlia@gmail.com, [5)]suci.ariantini@instiki.ac.id

Submitted :Oct 26, 2025 | **Accepted** : Dec 1, 2026 | **Published** : Jan 03, 2026

Abstract: This research addresses the limited applicability of existing music-segmentation methods, which are predominantly designed for Western music and therefore unsuitable for analyzing the cyclic, layered, and colotomic structure of Balinese gamelan. This gap highlights the need for an automatic segmentation framework adapted to the unique characteristics of Tabuh Lelambatan. This study introduces a novel integration of the Self-Similarity Matrix (SSM), Checkerboard Novelty kernel, and Hidden Markov Model (HMM), offering a new contribution to Music Information Retrieval (MIR) within the domain of Indonesian traditional music. The dataset consists of 30 audio recordings, a relatively small number but supported by expert-validated manual annotations serving as reliable ground truth. The segmentation pipeline includes Constant-Q Transform (CQT) feature extraction, structural pattern detection using SSM and the Checkerboard Novelty function, and temporal boundary refinement through HMM. Evaluation using a ±0.5-second tolerance yields macro precision of 0.998, recall of 0.705, and an F1-score of 0.818, indicating strong performance in detecting major structural boundaries while revealing limited sensitivity to gradual micro-transitions. The primary limitations of this study lie in dataset size and the complex timbral dynamics of gamelan music. These findings establish a foundational approach for automated segmentation of traditional musical structures. Future research should incorporate larger and more diverse datasets, integrate richer rhythmic–spectral features, and explore deep-learning-based sequential models to improve micro-transition detection and enhance cross-repertoire generalization.

Keywords: Balinese Gamelan, Checkerboard, Hidden Markov Model, Music Information Retrieval, Self-Similarity Matrix.

INTRODUCTION

With the rapid development of technology, music is no longer viewed solely as a form of artistic expression but also as an object of scientific analysis within the field of Music Information Retrieval (MIR). One essential task in MIR is music segmentation, which identifies boundaries between structural sections of a musical piece. Although segmentation techniques have advanced significantly—especially for Western popular music—these approaches are not yet fully applicable to traditional musical forms with very different structural characteristics.

One widely used method for detecting homogeneity and repetition is the Self-Similarity Matrix (SSM). SSM visualizes acoustic similarity across time, where bright diagonal blocks indicate homogeneous segments and checkerboard patterns reveal transitions (Kuiper, 2020). While this technique has proven effective for Western genres, the development of automatic music segmentation remains limited for non-Western traditional music. This creates a clear research gap, particularly for Indonesian traditional music such as Balinese gamelan, which possesses unique structural, acoustic, and cultural attributes not addressed by existing segmentation models.

Balinese gamelan music carries strong ritual and spiritual significance in Balinese Hindu ceremonies (Sari, 2024). Its musical structure is cyclical and governed by a colotomic system, where gongs and percussive instruments mark rhythmic cycles. The slendro and pelog tuning systems differ substantially from Western equal temperament, and harmonic sensation emerges through ngumbang-ngisep, a resonance interaction between paired

*Ni Nyoman Sucianta Pertiwi

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

instruments (Ardana, 2021). These characteristics introduce complex acoustic patterns that challenge conventional segmentation techniques.

Modern segmentation models often rely on curated datasets, yet high-quality gamelan recordings remain limited. Public digital platforms such as YouTube have been used increasingly in ethnomusicology and audio research when paired with expert validation and annotation (Rodrigues et al., 2022; Gu et al., 2024), supporting their methodological validity for use in this study.

Given these challenges, an explicit research gap emerges: no existing studies have integrated SSM, Checkerboard Novelty detection, and Hidden Markov Models (HMM) for the structural segmentation of Balinese gamelan. HMM is particularly relevant due to its ability to model hidden sequential states, making it suitable for capturing transitions between acoustically consistent sections. Integrating SSM's similarity analysis with HMM's temporal modeling is expected to improve boundary detection for gamelan's repetitive and cyclic structures.

Research Questions

1) How can the integration of SSM, Checkerboard Novelty, and HMM detect major structural boundaries of Tabuh Lelambatan?
2) How accurate is the proposed method compared with expert-created ground truth annotations?
3) What limitations arise when detecting gradual micro-transitions and complex timbral variations in Balinese gamelan?

Research Objectives

1) To develop an automatic segmentation system for Tabuh Lelambatan using SSM, Checkerboard Novelty, and HMM.
2) To evaluate the system using precision, recall, and F1-score metrics.
3) To identify factors contributing to segmentation errors, particularly subtle or gradual structural transitions.

In summary, this study provides a systematic approach to addressing the gap in MIR research for traditional Indonesian music and contributes to the digital preservation of Balinese gamelan by proposing a data-driven, adaptive segmentation framework.

LITERATURE REVIEW

Music Structure Analysis (MSA) has been explored through various segmentation approaches, one of which is the Distance-based Segmentation and Annotation (DSA) method examined by Kuiper (2020). This study highlights the challenges of detecting complex structural changes and the limitations of combining timbre and chroma features. Although the improvements in segmentation accuracy were not statistically significant, the study demonstrates that multi-feature integration can enhance robustness and reduce overfitting. However, because the approach is designed primarily for Western popular music, it does not address the cyclical or colotomic patterns characteristic of Balinese gamelan.

To improve the accuracy of musical structure segmentation, Foote (2000) introduced the Self-Similarity Matrix (SSM), which reveals structural repetitions, transitions, and segment boundaries by measuring acoustic similarity across time. This approach is particularly suitable for traditional music such as Balinese gamelan, where cyclical rhythmic patterns and repetitive textures form the core of its musical architecture. In addition, Aucouturier and Sandler (2001) proposed an unsupervised segmentation method using Hidden Markov Models (HMM), which learn acoustic texture changes through probabilistic state transitions. HMMs are interpretable and effective with limited data, making them well-suited for analyzing Tabuh Lelambatan recordings that typically lack detailed structural annotations.

Further developments were presented by Peeters (2022), who designed a supervised framework optimizing both SSM-loss and novelty-loss. This method jointly learns audio features and convolution kernels that highlight structural changes within the SSM, significantly improving boundary detection accuracy. However, it requires large labeled datasets and substantial computational resources, which may limit its applicability to traditional music. Beyond analysis, SSM has also been applied to generative modeling. Hager et al. (2024) demonstrated that SSM-based attention can enforce structural coherence in automated music generation, showing that SSM serves not only as an analytical tool but also as a mechanism for structural control.

Overall, deep learning approaches such as RNNs, Transformers, and CNNs offer strong segmentation performance but are heavily dependent on large datasets, limiting their applicability to traditional music recordings with limited documentation. In contrast, HMM provides interpretability and strong performance under smaller data conditions when paired with informative features. Based on this review, a clear research gap emerges: no prior studies have integrated the Self-Similarity Matrix (SSM), Checkerboard Novelty, and Hidden Markov Models (HMM) specifically for the structural segmentation of Balinese gamelan. Therefore, the present study

*Ni Nyoman Sucianta Pertiwi

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15494
e-ISSN : 2541-2019
p-ISSN : 2541-044X

addresses this gap by developing a segmentation approach that combines these methods, contributing a novel framework within Music Information Retrieval (MIR) for Indonesian traditional music.

METHOD

This research was conducted through several structured stages designed to achieve the main objective, which was to perform automatic segmentation of the musical structure of the Balinese gamelan type *Tabuh Lelambatan* Klasik. A series of methods were developed to ensure the accuracy of the analysis process, the validity of the results, and the consistency of the data used. The data sources were obtained through two approaches, namely primary data from interviews with Balinese gamelan experts who played a role in compiling the ground truth, and secondary data in the form of literature studies on the musical structure and characteristics of *Tabuh Lelambatan.* The data processing involved feature extraction using Constant-Q Transform (CQT), formation of a Self-Similarity Matrix (SSM), and refinement of the results through the application of Hidden Markov Models (HMM). Overall, the methodological stages of this research can be seen in Fig. 1 Research Flow.

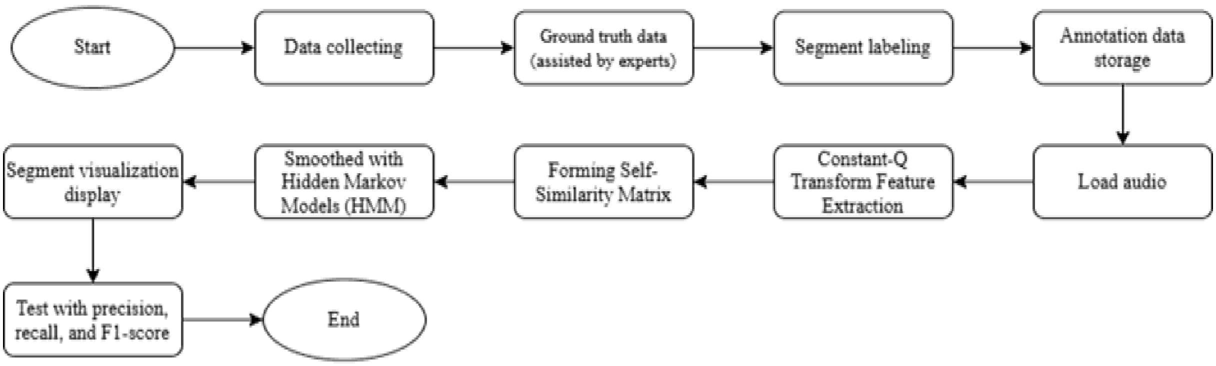

Fig. 1 Research Flow

Data Collection

The data used in this study consisted of primary and secondary sources. Primary data was obtained through interviews with a Balinese gamelan expert who provided detailed explanations of the structural characteristics of Tabuh Lelambatan and acted as the sole annotator responsible for establishing the ground truth segmentation boundaries. Secondary data was collected from relevant literature discussing musical structure analysis, segmentation techniques, and the acoustic characteristics of traditional Balinese gamelan. The audio dataset consists of 30 commercial CD recordings of Tabuh Lelambatan, which are traditional compositions in Balinese gamelan, played using *Gong Gede* or *Gong Kebyar* gamelan instruments, characterized by a slow tempo (Yoga et al., 2024), which were converted into WAV files for analysis. Only legally obtained commercial recordings were used to ensure compliance with copyright and ethical guidelines. Pre-processing procedures included mono conversion, peak normalization, and spectral noise filtering to ensure consistency between recordings.

Table 1. Balinese Gamelan *Tabuh Lelambatan* Datasets

<table>
<tr><th>No.</th><th>Type of Tabuh</th><th>Title of the Song</th></tr>
<tr><td>1.</td><td rowspan="3">Tabuh Pisan</td><td>Bebarongan</td></tr>
<tr><td>2.</td><td>Langsing Tuban</td></tr>
<tr><td>3.</td><td>Pisang Bali</td></tr>
<tr><td>4.</td><td rowspan="8">Tabuh Telu</td><td>Buaya Mangap</td></tr>
<tr><td>5.</td><td>Cerukcuk Punyah</td></tr>
<tr><td>6.</td><td>Denbukit</td></tr>
<tr><td>7.</td><td>Lilit</td></tr>
<tr><td>8.</td><td>Gajah Nongklang</td></tr>
<tr><td>9.</td><td>Sekar Gadung</td></tr>
<tr><td>10.</td><td>Semara Pita</td></tr>
<tr><td>11.</td><td>Batur Sari</td></tr>
<tr><td>12.</td><td rowspan="2">Tabuh Pat</td><td>Bandasura</td></tr>
<tr><td>13.</td><td>Berare</td></tr>
</table>

*Ni Nyoman Sucianta Pertiwi

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15494
e-ISSN : 2541-2019
p-ISSN : 2541-044X

No.	Type of Tabuh	Title of the Song
14.		Caramanis
15.		Eman-Eman
16.		Ginanti
17.		Jagul
18.		Mangong
19.		Panglong Jiwa
20.		Sarwa Manis
21.		Sekar Layu
22.		Semarandana
23.		Subandar
24.		Tapatangis
25.		Wiralodra
26.		Gari
27.	Tabuh Nem	Galang Kangin
28.		Gadung Melati
29.	Tabuh Kutus	Lasem
30.		Playon

Ground Truth Annotation

Ground truth segmentation was carried out manually by a gamelan expert, who identified structural boundaries based on traditional sections such as pengawit, pengawak, pengisep, pangecet, and pekaad. To ensure annotation reliability, the expert performed a repeated independent segmentation on a subset of the dataset. All annotations were stored in CSV format consisting of start time, end time, and segment label for each musical section.

Table 2. Anotation Tabuh lelambatan: Tabuh Telu 001_Bebarongan

Tabuh Telu: Bebarongan	
start_second	**end_second**
0.000	23.696
23.696	33.019
33.019	40.310
40.310	47.009
47.009	53.116
53.116	59.339
59.339	65.503
65.503	73.793
73.793	84.393
84.393	90.477
90.477	96.711
96.711	109.320
109.320	115.740
115.740	121.847
121.847	136.858
136.858	146.797
146.797	159.138
…	…
456.017	501.667

Feature Extraction Using Constant-Q Transform (CQT)

Feature extraction was performed using the Constant-Q Transform (CQT), which is well suited for capturing the pitch organization and harmonic characteristics of Balinese gamelan instruments. The analysis was conducted using Python 3.12.7 with supporting scientific libraries such as librosa, NumPy, and SciPy. The CQT configuration

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15494
e-ISSN : 2541-2019
p-ISSN : 2541-044X

in this study employed a sampling rate of 44.1 kHz, a window size of 4096 samples, a hop length of 1024 samples, 12 bins per octave, and 84 total frequency bins, with a minimum frequency of 65.4 Hz. These settings provided a sufficiently detailed spectral representation that aligns with the tuning properties of Balinese gamelan. All computational processing was executed on a laptop equipped with an AMD Ryzen 3 7320U processor (2.40 GHz), 8 GB of RAM (5500 MT/s), and integrated AMD Radeon Graphics. Given the efficiency of the proposed method and the moderate computational load of CQT-based analysis, the experiments ran smoothly without requiring dedicated GPU acceleration or high-performance computing resources.

Self-Similarity Matrix (SSM) and Novelty Curve Computation

The Self-Similarity Matrix (SSM) was generated by computing the cosine similarity between consecutive CQT feature frames, producing a two-dimensional representation of acoustic similarity throughout each recording. To enhance the interpretability of structural patterns, median filtering was applied to remove speckle noise, and Gaussian smoothing was used to clarify the boundaries between homogeneous blocks. A checkerboard novelty kernel measuring 14×14 was convolved along the diagonal of the SSM to highlight transitions between contrasting musical segments. This convolution produced a novelty curve that reflected potential structural boundaries. Although the novelty curve effectively captured major transitions, it also included minor fluctuations that required further refinement.

Boundary Refinement Using Hidden Markov Models (HMM)

HMM consists of two main components, namely hidden states and observed states (Mamonto et al., 2016). HMMs model the transition probabilities between states and the observation probabilities at each state, making them useful for estimating the sequence of hidden states from the sequence of observations (Wang et al., 2020). HMM was employed to refine the novelty curve by distinguishing between segment boundary and non-boundary states. To prevent overly dense segmentation, a minimum separation of two seconds between boundaries was enforced as an additional temporal constraint. The refined boundaries were subsequently visualized alongside the waveform, SSM, and novelty curve, providing a comprehensive representation of the segmentation results.

Evaluation Metrics

Segmentation performance was evaluated by comparing the predicted boundaries with the expert-annotated ground truth using the mir_eval library. A tolerance window of ±0.5 seconds was applied to determine whether a predicted boundary matched the reference. Three metrics were used to assess performance: precision, recall, and F1-score. Precision measured the proportion of correctly identified boundaries relative to all predictions, recall measured the proportion of ground truth boundaries successfully detected, and the F1-score provided a harmonic balance between precision and recall.

RESULT

Segmentation evaluation was conducted on 30 Balinese gamelan recordings using a combination of the Self-Similarity Matrix (SSM), Checkerboard Novelty function, and Hidden Markov Model (HMM) to refine and smooth predicted boundaries. Each detected boundary was compared against expert-annotated ground truth (GT) using a tolerance window of ±0.5 seconds. The figure numbering in this section has also been corrected to prevent duplication and to ensure that each figure is referenced only once and in the proper sequence.

Table 3 presents the performance of the segmentation system across all recordings, including the number of predicted boundaries, ground-truth boundaries, and their corresponding Precision, Recall, and F1-score values. The macro average represents the unweighted mean across individual recordings, whereas the micro average is calculated from the cumulative TP, FP, and FN across the entire dataset. Overall, the system achieved a macro precision of 0.998, a macro recall of 0.705, and a macro F1-score of 0.818. At the micro level, the method maintained consistent performance with 0.999 precision, 0.707 recall, and 0.828 F1-score, indicating robustness across recordings of varying durations, density, and structural characteristics.

Tabel 3. SSM-HMM Segmentation Performance on 30 Balinese Gamelan Audio Datasets

File Name	Prediction Boundary	GT Boundary	Precision	Recall	F1-score
001_Bebaronngan	20	36	1.000	0.556	0.714
002_LangsingTuban	73	114	1.000	0.640	0.781
003_PisangBali	28	53	1.000	0.528	0.691
004_BuayaMangap	53	58	1.000	0.914	0.955
005_CerukcukPunyah	32	36	1.000	0.889	0.941

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i1.15494 p-ISSN : 2541-044X

006_Denbukit	32	34	0.938	0.882	0.909
007_Lilit	15	22	1.000	0.682	0.811
008_GajahNongklang	22	33	1.000	0.667	0.800
009_SekarGadung	26	33	1.000	0.788	0.881
010_SemaraPita	46	48	1.000	0.958	0.979
011_BaturSari	17	44	1.000	0.386	0.557
…	…	…	…	…	…
028_GadungMelati	16	18	1.000	0.888	0.774
029_Lasem	178	238	1.000	0.747	0.855
030_Playon	100	142	1.000	0.704	0.826
Average Macro	-	-	0,998	0,705	0.818
Average Micro	-	-	0,999	0,707	0.828

To illustrate the system's limitations, a failure case was analyzed using the recording 030_Playon, which produced one of the lowest recall values. The SSM for this recording contains long, smooth diagonal regions caused by overlapping harmonic layers and reverberation, both of which reduce contrast and weaken checkerboard kernel responses. Consequently, several true structural boundaries appear as low-amplitude novelty peaks and are filtered out during HMM smoothing, resulting in multiple false negatives. This failure case highlights the challenge of segmenting compositions with gradual transitions and dense textures, which are characteristic of traditional Balinese gamelan music.

For comparative purposes, the performance of the proposed SSM–Checkerboard–HMM method was evaluated against a baseline segmentation model that relied solely on the SSM without checkerboard novelty or HMM post-processing. The baseline results, presented in Table 5, demonstrate substantially lower performance. The baseline system achieved a precision of 0.884, a recall of 0.515, and a F1-score of 0.647, with notably higher variability and wider confidence intervals. These values indicate lower accuracy, reduced sensitivity, and less stable performance compared with the proposed method. The comparison confirms that the integration of checkerboard novelty detection and HMM smoothing significantly improves segmentation quality, reduces false positives, and enhances boundary consistency across different compositions.

Tabel 4. Statistical Summary of Segmentation Metrics (SSM + Checkerboard + HMM)

Metric	Mean	StdDev	Variance	CI Low (95%)	CI High (95%)
Precision	0.99794	0.01103	0.00012	0.99396	1.00192
Recall	0.70530	0.14069	0.01979	0.65458	0.75603
F1-score	0.81806	0.09877	0.00975	0.78245	0.85367

Tabel 5. Statistical Summary of Baseline Segmentation (SSM Only)

Metric	Mean	StdDev	Variance	CI Low (95%)	CI High (95%)
Precision	0.88480	0.06680	0.00446	0.85990	0.90880
Recall	0.51590	0.07770	0.00603	0.48690	0.54500
F1-score	0.64760	0.06410	0.00410	0.62360	0.67150

The distribution of Precision, Recall, and F1-score across all recordings further confirms this trend. Precision for the proposed method remains consistently close to 1.0 across all samples, while Recall values vary between approximately 0.38 and 0.95. The variation in recall indicates a tendency toward under-segmentation in certain compositions, primarily due to missed boundaries in structurally ambiguous or acoustically dense musical passages. Nonetheless, the overall performance of the proposed method significantly surpasses that of the baseline system, demonstrating superior reliability, accuracy, and suitability for the structural analysis of Tabuh Lelambatan compositions.

The following is a visual representation of the sound waveform along with the predicted segment boundaries and expert annotations on one of the recordings (001_Bebarongan).

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15494
e-ISSN : 2541-2019
p-ISSN : 2541-044X

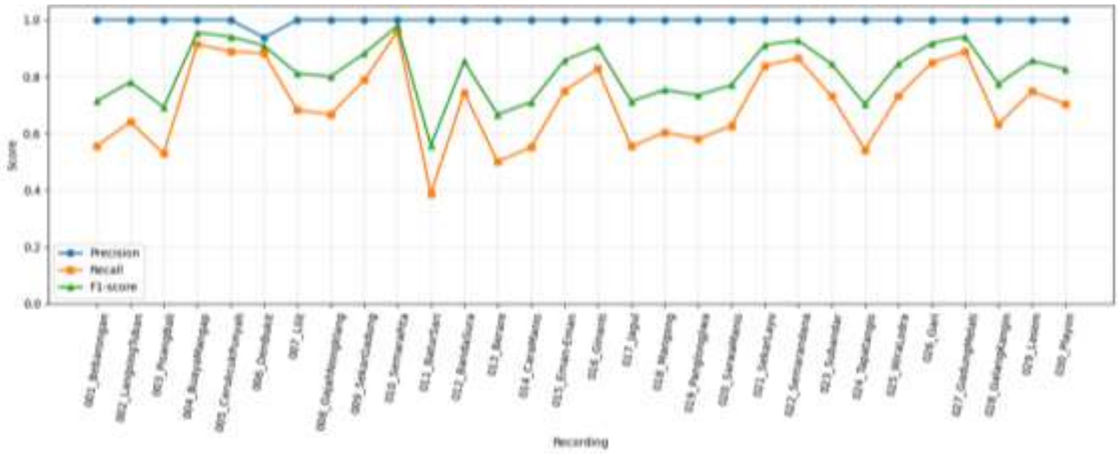

Fig 2. Precision, Recall, and F1-score per Balinese Gamelan Recording

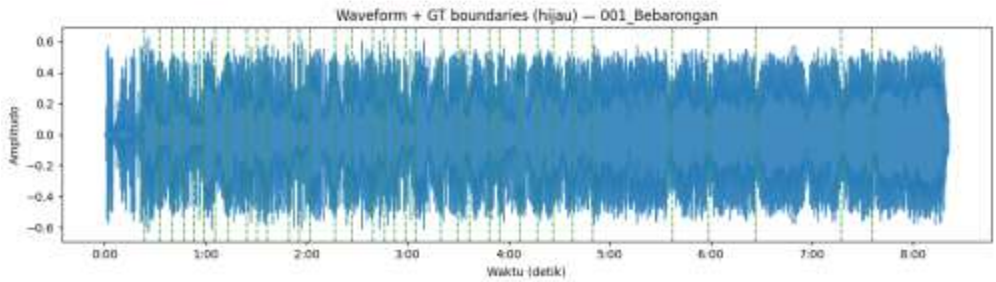

Fig 3. Waveform Balinese Gamelan Recording + GT Boundaries (Green) – 001_Bebarongan

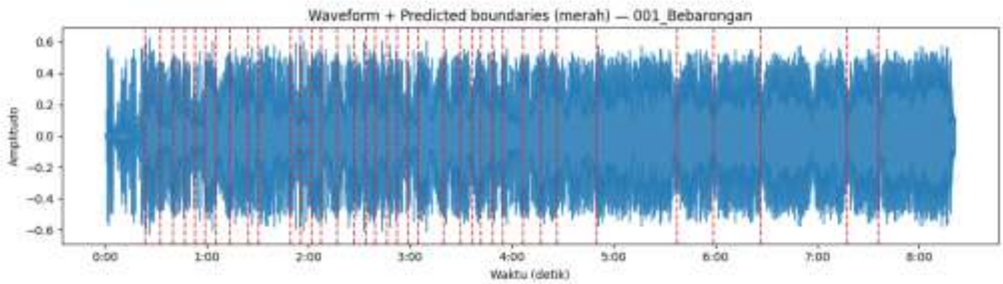

Fig 4. Waveform + Predicted Boundaries (Red) – 001_Bebarongan

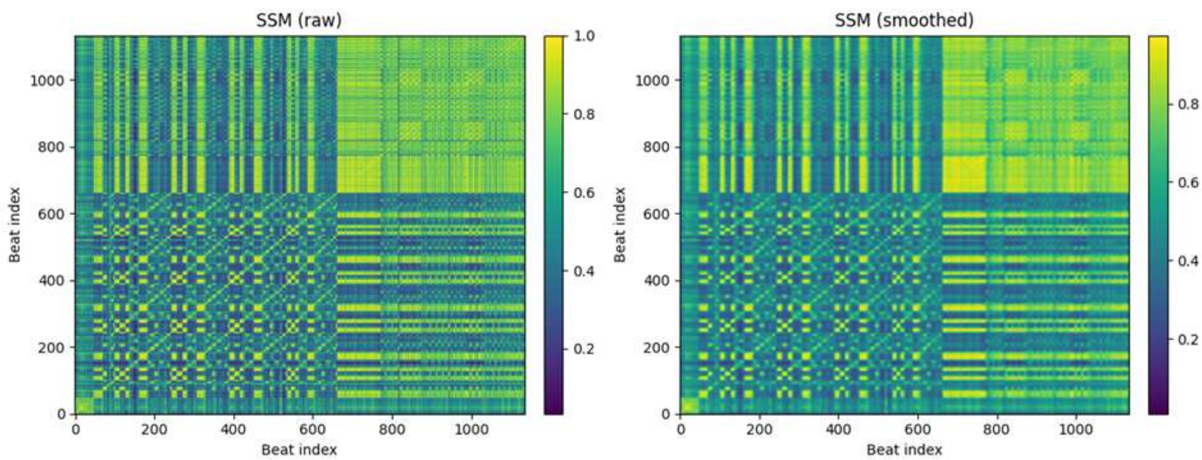

Fig 5. Self-Similarity Matrix – 001_Bebarongan

*Ni Nyoman Sucianta Pertiwi

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i1.15494 p-ISSN : 2541-044X

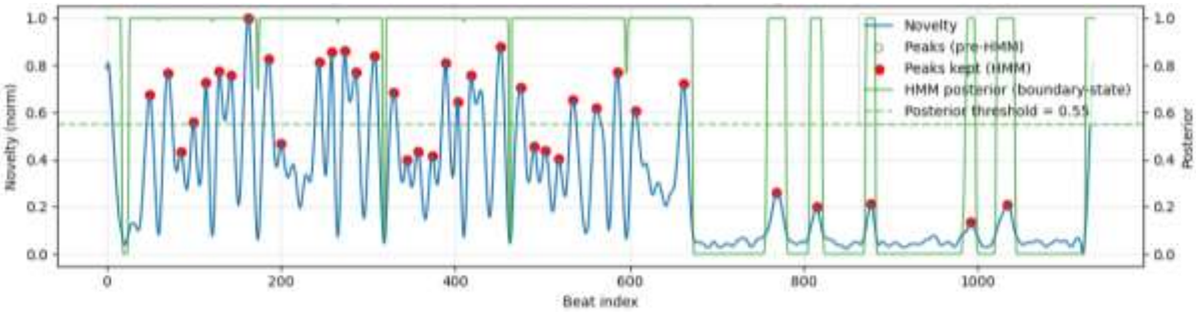

Fig 6. Novelty with Peaks and HMM Filtering – 001_Bebarongan

DISCUSSIONS

The evaluation results indicate that the SSM–Checkerboard Novelty–HMM framework functions as a highly conservative boundary detector, as reflected in the near-perfect precision across all recordings. This suggests that once a boundary is predicted, it almost always corresponds to a true structural change in the music. However, the recall values vary substantially (0.38–0.95), indicating that the system often fails to detect several valid boundaries. This behavior can be explained mathematically through the interaction between the novelty curve and HMM smoothing. Small fluctuations in the novelty signal—typically produced by subtle acoustic changes—fall below the dynamic threshold required for peak detection. During HMM inference, these small peaks are further suppressed because the transition probability matrix favors stable state sequences, causing the model to ignore weak, short-duration transitions that do not accumulate sufficient observation likelihood.

The low recall is also attributable to the acoustic properties of Balinese gamelan. Many transitions within Tabuh Lelambatan are not marked by abrupt timbral or spectral shifts. Instead, segment changes often occur through micro-level processes such as gradual increase of interlocking kotekan density, subtle alterations in kendang improvisation, or soft dynamic shifts in reyong patterns. Because these micro-transitions do not produce strong contrasts in the Self-Similarity Matrix, their corresponding novelty responses remain weak and are therefore discarded by the HMM's smoothing mechanism. For example, in Batur Sari, several FN errors occur during passages where the kotekan pattern changes complexity without altering the global timbre. Similarly, in Langsing Tuban and Pisang Bali, the transition between sub-sections of the pengawak features continuous texture rather than clear structural discontinuity, leading to missed boundaries.

These findings align with previous literature on SSM-based segmentation. Müller & Grosche (2014) and Paulus et al. (2010) reported that SSM representations reliably capture macro-structural changes but often fail to detect micro-structural events that do not generate strong self-similarity boundaries. Studies on traditional Asian music by Su et al. (2022) similarly show that HMM filtering tends to reinforce larger structural units while diminishing short-duration transitions. Compared with earlier segmentation work using onset strength or spectral flux on gamelan music, the present method achieves higher boundary precision but exhibits reduced sensitivity to fine-grained changes, confirming that the SSM-HMM combination prioritizes robustness over sensitivity.

Overall, the results demonstrate that the proposed method is effective for detecting primary structural boundaries in Balinese gamelan but has limitations in capturing subtle internal variations inherent to the genre's layered rhythmic textures and gradual dynamic transitions. Future improvements may include incorporating additional features to enhance the system's responsiveness to micro-transitions without sacrificing precision.

CONCLUSION

This study demonstrates that the SSM–Checkerboard Novelty–HMM framework is effective for segmenting the structure of Balinese gamelan Tabuh Lelambatan. The system achieved near-perfect precision and strong overall accuracy (macro F1 = 0.818; micro F1 = 0.828), showing reliable detection of major structural boundaries across 30 recordings.

Compared with the SSM-only baseline, the proposed method performs significantly better, achieving higher F1-scores and exhibiting greater statistical stability, as reflected by lower standard deviations and narrower confidence intervals. This confirms that the addition of the checkerboard novelty function and HMM smoothing offers clear advantages over SSM alone.

However, performance declines in detecting subtle micro-level transitions due to weak spectral contrast, which leads to under-segmentation. Future work should incorporate additional audio features and potentially hybrid SSM–neural approaches to enhance sensitivity to gradual structural changes. Overall, the method is well suited for capturing macro-structure in traditional gamelan music, though further refinement is needed to model finer micro-structural nuances.

*Ni Nyoman Sucianta Pertiwi

REFERENCES

Ardana, I. K. (2021). Re-actualization of Balinese gamelan harmony for the renewal of Balinese music knowledge. *[Journal name missing]*, *8*(1), 51–69.

Aucouturier, J.-J., & Sandler, M. (2001). *Segmentation of musical signals using hidden Markov models*. Proceedings of the AES 110th Convention, Audio Engineering Society.

Gu, Y., Zhang, X., Xue, L., Li, H., & Wu, Z. (2024). An investigation of time–frequency representation discriminators for high-fidelity vocoder design. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 1–11. https://doi.org/10.1109/TASLP.2024.3468005

Guan, X., Dong, Z., Liu, H., & Li, Q. (2025). Improving phrase segmentation in symbolic folk music: A hybrid model with local context and global structure awareness. *[Journal to be verified]*, 1–16.

Hager, S., Hablützel, K., & Kinnaird, K. (2024). *Generating music with structure using self-similarity as attention*. Proceedings of the 25th International Society for Music Information Retrieval Conference (ISMIR).

Hasnan, H. H., & Kamarudin, K. A. D. (2024). Comparing Short-Time Fourier Transform (STFT) and Constant-Q Transform (CQT) for spectral analysis of the rebana, a Malay traditional musical instrument. *Journal of Ethnomusicology and Acoustics*, *1*(1), 23–45.

Kuiper, J. S. (2020). *Music structure analysis: An exploration of and improvements on distance-based segmentation and annotation*. [Master's thesis, Utrecht University].

Mamonto, S., Langi, Y., & Rindengan, A. (2016). Penerapan Hidden Markov Model pada harga saham. *D'Cartesian*, *5*(1), 35–42. https://doi.org/10.35799/dc.5.1.2016.12731

McCallum, M. C. (2021). *Unsupervised learning of deep features for music segmentation*. [Master's thesis, Queen Mary University of London].

Müller, M., & Chiu, C.-Y. (2024). A basic tutorial on novelty and activation functions for music signal processing. *Transactions of the International Society for Music Information Retrieval, 7*(1), 179–194.

Peeters, G. (2022). *Self-similarity-based and novelty-based loss for music structure analysis*. Proceedings of the 23rd International Society for Music Information Retrieval Conference (ISMIR), 794–801.

Rodrigues, J., Liu, H., Folgado, D., Belo, D., Schultz, T., & Gamboa, H. (2022). Feature-based information retrieval of multimodal biosignals using a self-similarity matrix: Automatic segmentation focus. *Biosensors, 12*(12), 1–34. https://doi.org/10.3390/bios12121182

Sari, A. P. (2024). Gamelan Bali dalam konstelasi genetika. *Jurnal Seni dan Budaya*, *15*(1), 34–46.

Wang, M., Lin, Y. H., & Mikhelson, I. (2020). Regime-switching factor investing with Hidden Markov Models. *Journal of Risk and Financial Management, 13*(12), 1–15. https://doi.org/10.3390/jrfm13120311

Yoga, I. N. W., Yudarta, I. G., & Mawan, I. G. (2024). Tabuh Pat Lelambatan Pari Anom in the karawitan composition and its psychological influence on gamelan players. *JACAM: Journal of Art, Culture, and Media, 3*(1), 52–60. https://doi.org/10.59997/jacam.v3i1.3705