

AI-ASSISTED NOTATION SYSTEMS IN MUSIC PEDAGOGY

Abhijeet Panigra 1✉, Ganesh Korwar 2✉ , RPS Chauhan 3✉, Aswitha V. 4✉, Somanath Sahoo 5✉ , Pooja Nagargoje 6✉

¹ Assistant Professor, School of Business Management, Noida International University, Greater Noida, 203201, India

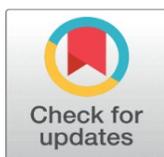
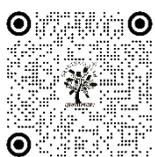
² Department of Mechanical Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, 411037, India

³ Department of Artificial Intelligence and Machine Learning, Shri Shankaracharya Institute of Professional Management and Technology, Raipur, Chhattisgarh, India

⁴ Assistant Professor, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, 600091, India

⁵ Associate Professor, School of Journalism and Mass Communication, AAFT University, Raipur, Chhattisgarh-492001, India

⁶ Researcher Connect Innovation and Impact Pvt. Ltd, Nagpur, Maharashtra, India



Received 16 September 2025

Accepted 19 December 2025

Published 17 February 2026

Corresponding Author

Abhijeet Panigra,

abhijeet.panigrahy@niu.edu.in

DOI

[10.29121/shodhkosh.v7.i1s.2026.7109](https://doi.org/10.29121/shodhkosh.v7.i1s.2026.7109)

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Copyright: © 2026 The Author(s). This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.



ABSTRACT

In this paper, an AI-assisted notation system is suggested to assist beginners, intermediate, and advanced musicians in the creation of scores with the help of intelligent score generation, a real-time experience, and an intelligent, pedagogically adjusted approach. The paper discusses the drawbacks of traditional teaching notation such as a slow response, fixed display, and heavy workload among beginners. The proposed method uses a combination of signal processing with pitch, rhythm, and timbre determination and deep learning models such as convolutional and recurrent models and transformer-based models to attain accurate results in automatic music transcription and expressive analysis. MIDI and MusicXML are symbolic representations that are used to ensure educational readability and interoperability. Training of the models is done using multi-instrument data using stratified validation in various levels of learner proficiency. Assessment is based on the accuracy of transcription, precision of rhythm, comprehensibility of the notation, and pedagogical convenience using structured tasks of learners. The experimental data prove that an AI-assisted system has a mean pitch accuracy of 91.6, a rhythmic accuracy of 88.3, and a learning error based on the notation is much higher (27 percent in comparison with traditional training). Students with adaptive notation have an increase in their speed of sight-reading of 22 percent and an increase in their engagement scores, especially with beginners.

Keywords: AI-Assisted Notation, Music Pedagogy, Automatic Music Transcription, Symbolic Music Representation, Adaptive Learning Systems

1. INTRODUCTION

Music notation has been traditionally the major medium of capturing, transmitting, and teaching musical ideas. Being a symbolic form of sound, notation allows learners to encode pitch, rhythm, dynamics, articulation and structure into visual forms which encode information about sound through the use of symbols. The technical skill, theoretical

knowledge, and the long-term growth of music are closely interconnected in formal music education when it comes to reading and interpreting the notation. Nevertheless, the conventional methods of teaching notation usually incorporate the use of static scores, slow response by the instructor and the same pace of pedagogy which might not sufficiently meet the various learning requirements, cognitive burdens and skill development of modern learners. As the use of digital technologies in education continues to grow, music pedagogy has undergone a slow but steady change in the learning models of complete instructor control to technology-supported and student-focused ones [Cui \(2023\)](#). The use of computers in teaching music also added to the interactive activities, computer playback and basic music analysis, but most of the computer-based music teaching systems lack flexibility and musical comprehension. Specifically, traditional digital notation tools are usually passive editors or display interfaces, which provide little understanding of how students perform, their areas of weakness or in what ways notation itself would support learning. New possibilities in the development of the role of notation in music education are offered by recent advances in artificial intelligence. In AI-assisted notation, notation systems are no longer passive representations of musical data, but act upon musical input, create symbolic scores on their own and change their visual complexity according to the proficiency of the learner [Chen \(2022\)](#). Such systems can close the performance-notation gap because, via signal processing, machine learning and representation of symbolic music, the student can observe immediate, intelligible displays of what he or she is playing or singing. This live connection between act and symbol can greatly increase music perception, perception of errors and conceptual comprehension [Zhang \(2023\)](#).

One of the most important reasons why AI-assisted notation should be part of pedagogy is the difficulty that beginners and non-conventional students have. Novice students tend to feel overly cognitive as they have to encounter thick scores that encode different musical dimensions at the same time. However, in difference, more progressive scholars might need descriptive annotations and critical comments in order to perfect interpretation. One, fixed notation format is thus pedagogically inefficient on skill levels. This problem is solvable through AI-based systems that will produce adaptive notation that adds complexity over time, optimal features, and visual representation in line with the instructional objectives [Zheng \(2024\)](#). Additionally, the learning of music is performance-based in nature, but conventional notation evaluation of learning often isolates practice and evaluation. Students usually get feedback in performance that is usually distorted with subjective instructor judgment. The AI-assisted systems of notation allow correcting the performance based on the feedback data by comparing the performance of the student to the reference models on a continuous basis and pointing out the pitch deviation, rhythmic discrepancies, and expressive discrepancies within the notated score [Chin and Xia \(2022\)](#). This forms a closed feedback loop where learners get to identify their errors immediately in relation to certain symbolic elements, which makes them self-regulate their learning and the efficiency of their practice. Scalable and inclusive music education is another application of AI-assisted notation, which has an institutional angle.

2. LITERATURE REVIEW

2.1. CONVENTIONAL MUSIC NOTATION AND TEACHING FRAMEWORKS

Traditional music notation has traditionally served as the language of the formal study of music in which it serves as a standardized system to notate pitch, rhythm, meter, dynamics, articulation and form. Institutional pedagogy, in its turn, has been dominated by western staff notation, with its models of conservatoires, graded syllabus and examination-focused teaching. The conventional frameworks focus on score reading, imitation, repetition, and instructor-guided correction with the learners learning the notation by being taught and explained by a teacher [Knapp et al. \(2023\)](#). The method encourages musical literacy and the continuation of repertoire over a long period of time but usually presupposes a homogenous learning rate and prior acquaintance. Notation is usually presented pedagogically, starting with simple values of rhythm and pitch placement, and moving on to more complex forms of harmony and expressiveness. Whereas they may be effective in structured curricula, these frameworks are more likely to give an advantage to the visual-symbolic interpretation, at the expense of those learners who are aural or kinesthetic thinkers [Wan \(2024\)](#). Further, the feedback on a traditional environment is usually asynchronous and provided at the end of performance and mediated by the interpretation of the teacher, which can restrict the possibility of correcting errors quickly. The other weakness is the inability of traditional scores to change. After being submitted or fixed, notation is not sensitive to the errors or level of proficiency and contextual challenge that learners may have. The novices can experience a problem of cognitive overload when the decoding of different symbolic dimensions occurs in parallel, and the more

experienced learners can experience a lack of expressiveness when they need to analyze the nuances of expressiveness by means of conventional notation [Hou \(2024\)](#).

2.2. COMPUTER-ASSISTED MUSIC INSTRUCTION (CAMI) MODELS

Computer-aided music teaching became one of the first efforts at improving traditional pedagogy with interactive digital media. The CAMI systems are commonly combined with software-related tutorials, visual displays, playback features, and simple evaluation packages in order to facilitate the learning of music outside or during classroom instruction. First models referred to drills training exercises in rhythm, pitch recognition, sight-reading, music theory and provide immediate feedback, which was rule-based. These systems enhanced learner interaction and precision of practice, though they were weak in terms of musical intelligence and context awareness [Cao \(2022\)](#). Later models of CAMI also added multimedia features like synchronized audio-notation display, MIDI input and graphical performance indicators. Learners had the ability to see timing changes, wrong notes or changes in tempo, and thus could more objectively self-assess than by practice alone. Nonetheless, the majority of CAMI systems were based on predetermined thresholds and template matching as opposed to adaptive learning systems. Consequently, feedback was usually generic and unresponsive to personal development as well as expressiveness [Solanskyi et al. \(2024\)](#). Pedagogically, CAMI systems helped to achieve learner independence and scalability especially in distance learning situations. They favored drilling and testing but seldom used dynamism in the instructional content [Gourikeremath and Hiremath \(2025\)](#).

2.3. AUTOMATIC MUSIC TRANSCRIPTION (AMT) AND SYMBOLIC MUSIC AI

Music transcription Music notation Automatic music transcription is a field of important research in audio signal processing that connects audio signal processing to symbolic representation of music. AMT tries to encode unstructured audio performances into notation estimated of parameters pitch, onset, duration and occasionally expressive properties. The early AMT methods were dependent on digital signal processing tools, such as the Fourier analysis, spectral peak search, and rule of thumb sets. Although these techniques worked well with monophonic melodies, they did not perform well on polyphony, timbral variation and noisy performance environments [Zhang \(2023\)](#). Machine learning was also enhanced and made AMT more accurate, especially using deep learning architectures, including convolutional and recurrent neural networks. These models directly learn the complex time and spectral patterns on the basis of the data, thus being able to locate pitch information and rhythmic syncing as well as distinguish between instruments in a more robust fashion. Evolution of AI-assisted music notation, transcription, and pedagogy systems can be found in [Table 1](#). Transformer based models also provided better models of long range dependency, which facilitated transcription of expressive timing and structural coherence. Similar to AMT, symbolic music AI is concerned with representation, manipulation and generation of musical structures by using formats such as MIDI and MusicXML.

Table 1

Table 1 Related Work on AI-Assisted Notation, Music Transcription, and Pedagogical Systems					
Core Focus Area	Data Type Used	Algorithmic Method	Pedagogical Level	Key Contribution	Limitations
Digital music tutoring	MIDI, audio	Rule-based + DSP	Beginner	Early CAMI-based feedback	No adaptive learning
Automatic music transcription	Audio	CNN	General	Improved pitch detection	Weak rhythm modeling
Score-following systems Konecki (2023)	Audio + score	HMM	Intermediate	Real-time alignment	Limited expressivity
Piano transcription	Polyphonic audio	CNN + RNN	Advanced	Polyphonic accuracy gain	Low readability
Music education software	MIDI	CAMI heuristics	Beginner	Interactive learning	Static notation
End-to-end AMT Akgun and Greenhow (2022)	Audio	Transformer	General	Long-range dependency modeling	Pedagogy not addressed
Symbolic music learning	MIDI	LSTM	Intermediate	Sequence modeling	No visual adaptation

Expressive performance analysis	Audio	CNN + Attention	Advanced	Expressive feature extraction	Not learner-centric
Educational transcription tools	Audio	Hybrid DL	Beginner-Intermediate	Improved notation clarity	Limited personalization
Real-time music feedback González-González (2023)	Audio	Deep RNN	Beginner	Immediate feedback loop	High computational cost
Adaptive learning systems	MIDI	Reinforcement Learning	Intermediate	Learner adaptation	Limited musical scope
AI music pedagogy	Audio + MIDI	CNN + Transformer	Multi-level	Pedagogy-aware feedback	Dataset constrained

3. TECHNICAL FOUNDATIONS OF AI-ASSISTED NOTATION

3.1. SIGNAL PROCESSING FOR PITCH, RHYTHM, AND TIMBRE EXTRACTION

Extracting pitch is one of the main tasks, which is intended to determine the essential frequency of a musical signal in time. The classical methods of autocorrelation, cepstral and the short-time Fourier transform are commonly employed to estimate the pitch contours, especially monophonic sources. More complex methods, such as harmonic summation and probabilistic pitch tracking enhance the strength in the case of noise, vibrato, and expressive intonation changes, which are standard in pedagogical performances. The tasks of rhythm and temporal structure extraction aim at identifying the note onsets, duration and metrical alignment. Detectors Onset detectors are usually based on spectral flux or energy envelope analysis or phase deviation to detect transient events. Fig. 1 demonstrates the obtained pitch, rhythm, and timbre features that allow performing musical analysis. Rhythmic extraction Rhythmic extraction is also necessary because errors in timing or beats can severely diminish pedagogical clarity in the case of readability of the notation.

Figure 1

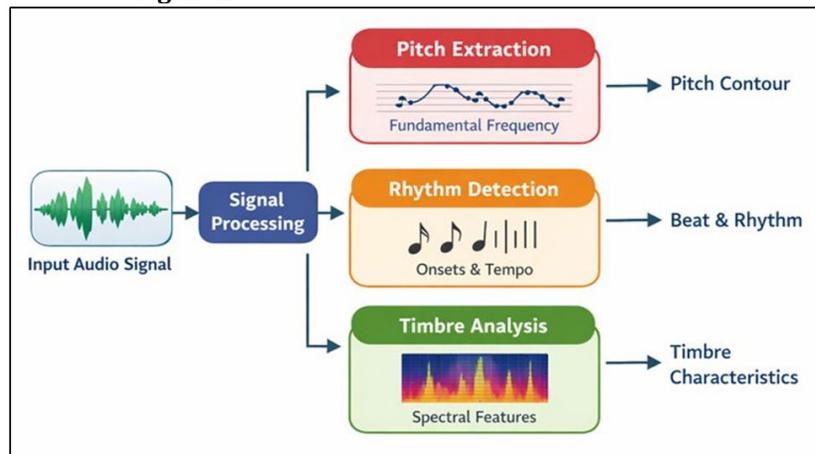


Figure 1 Feature Extraction Pipeline for Pitch, Rhythm, and Timbre Analysis

To the learners, rhythmic feedback in terms of visualization helps to develop timing awareness and metrical comprehension. Timbre extraction is the process that deals with spectral and feeling features of instruments and the playing method. Characteristics of the timbral variation are often quantified in Mel-frequency cepstral coefficients, spectral centroid, bandwidth and harmonic-to-noise ratio. Timbre analysis applied in an educational context allows the transcription of instruments in question, detection of articulation, and measurement of performance quality.

3.2. DEEP LEARNING MODELS

1) CNNs

Convolutional neural network has become an essential part of music analysis systems because it is able to learn spatially localized patterns based on time frequency representations of audio signals. The CNNs are generally used in AI-assisted notation systems, with spectrograms, constant-q transforms, or Mel-frequency representations being the usual

inputs, and regarded as two-dimensional images. CNNs automatically extract features that are related to the harmonic structure, onset of notes, pitch salience and textural components of the timbre. The feature minimizes the use of handcrafted characteristics and enhances the strength of different instruments and situations during recording. The CNN-based models can be effectively used with pitch and onset recognition, in which the local spectral patterns can be very helpful. CNNs are used in pedagogy to identify events at the note level with high accuracy based on student performances, which underlie the transcription on reliable grounds. CNNs are however, mostly good at capturing local context and might not be able to handle the long term temporal connections like phrasing or expressive timing.

2) Recurrent Neural Networks (RNNs)

RNNs are implemented to learn sequential data, and thus they are effective in learning temporal relationships that prevail in music. Other variants like long short-term memory networks and gated recurrent units overcome the vanishing gradient problem allowing the learning of long-range correspondence between musical phrases. RNNs are used in AI-assisted notation systems to predict continuity in pitch, rhythm and expressive timing at different points in time by taking sequences of audio features extracted. RNNs prove especially useful in modeling the rhythm, beat tracking and the duration of notes as the models retain an internal state which captures the previous musical context. This time consciousness enables the system to differentiate between performance mistakes that are isolated in time and those that are systematic time violations that is pedagogically useful.

3) Transformers

Transformers constitute a new paradigm shift in the analysis of music because they substitute it with repetition with self-attention mechanisms modeling the global dependencies of entire sequences. In AI-assisted notation systems, musical input is processed by transformers considering note-note relationships, note-note rhythm and other expressive features without regard to the temporal distance. This attribute is useful, especially in preserving long-range musical structure (i.e., repetitive motives, boundaries of phrases, and metrical consistency). In contrast to RNNs, transformers simultaneously process sequences, which is very efficient and scaled to a great extent. This renders them appropriate in real-time or near-real-time transcription and analysis in the learning context. Attention weights are also somewhat interpretable to enable educators and learners to learn what musical events contribute to transcription or feedback decisions. Transformers are particularly good at sequence-to-sequence tasks in symbolic music AI, including audio features to notation something to refine a preliminary transcription.

3.3. SYMBOLIC REPRESENTATION FORMATS (MIDI, MUSICXML, ABC)

The representation formats of music notation are increasingly becoming important in AI-assisted notation systems to enable the synthesis, analysis, and transformation of musical information into machine-readable formats to support pedagogical applications. MIDI is one of the most popular formats, which interprets music in the form of a series of performance-oriented events such as note onsets, durations, velocities, and control changes. Its small size and live-time compatibility render MIDI quite appropriate in recording the student performances and aid with instant feedback. Nevertheless, MIDI has no explicit notation semantics, including beaming, articulation symbols, and layout information, and is therefore not easily readable by educational notation purposes. MusicXML attempts to overcome these limitations through the provision of a notation-based representation that explicitly represents the staff structure, time signatures, signature of key, dynamics, articulations and engraving. Due to this, MusicXML has been universal in educational software and notation editors and allows scores to be high-quality and readable by humans. MusicXML can also be used in AI-assisted pedagogy to bridge the gap between automatic transcription output and a visually understandable notation that has been optimized to the abilities of the learner.

4. METHODOLOGY

4.1. MODEL TRAINING AND VALIDATION DESIGN

The model training and validation model is set in a way that allows it to be robust, generalized, and relevant to pedagogy of the AI-assisted notation system. Multi instrument data of monophonic and polyphonic performances of different tempos, dynamics, and expressive styles are gathered as audio inputs. Preprocessing involves the process of reducing noise, normalization and time-frequency transformation to produce standardized input representations. Learning pipeline has a modular structure where the feature extraction, transcription and symbolic refinement are

trained in a sequential or end-to-end style depending on the complexity of a task. Supervised learning is used in training where annotated ground-truth scores are mapped on performance audio. Tempo scaling, pitch shifting, and timbral variation are data augmentation techniques that are used to reduce the bias in the dataset and enhance the robustness of the model. Validation is done on a stratified split basis so that there is a balance in terms of instruments, difficulty level and profiles of learners. The stability is measured by cross-validation to eliminate overfitting.

4.2. USER GROUPS: BEGINNER, INTERMEDIATE, ADVANCED LEARNERS

The methodology clearly separates the user into beginner, intermediate and advanced learner level to assess the adaptability and instructional relevance according to the levels of proficiency. The novices are inadequate in terms of notation literacy, slower sight-reading, and more sensitive to mental overload. In this group, the system focuses on simplified notation, less density of symbols, visual representation of rhythm, pitch, and is tolerant to expressive variability. The beginner evaluation training is based on the monophonic melodies, constant tempo, and simple rhythmic patterns. Middle level learners have the basic skills of notation and better technical control but might lack more complicated rhythms, polyphony or the ability to use expression marks. In the case of this group, the system brings in a richer symbolic detail, such as articulation, dynamics, and rhythmic subdivisions. Adaptive feedback helps in pointing out structural mistakes and promoting the refinement of the interpretations. The advanced students are fluent in notation and performance, they need high accuracy in transcription and expression. This category is evaluated by polyphonic textures, tempo variations and style interpretation. The system produces high density analytically rich notation and fine-grained feedback that is professionally oriented. This kind of segmentation of users makes the methodology critical to measuring performance improvement in relation to pedagogical requirements and not homogenous technical standards. The provided learner-centered stratification will allow a systematic study of AI-assisted notation facilitation of progression, scaffolding and mastery by stages in education.

4.3. EVALUATION METRICS: ACCURACY, READABILITY, PEDAGOGICAL USABILITY

The metrics of evaluation will be constructed to represent the technical performance of the AI-assisted notation system as well as the effectiveness of the education system. The accuracy measures are used to measure transcription fidelity, i.e. pitch correctness, onset alignment, rhythmic correctness and note duration overlap against reference scores. These measures can be used to offer objective standards of the performance of the models in various musical situations and in different learning levels. Readability measurements are used to determine the readability and interpretability of formed notation. Among these, there are the density of symbols, the rhythmic conformity, the regularity of the staff structure, and the clearness of error visualization. Visual comprehensibility is assessed with the use of quantitative measurements including note collision rate and average symbol spacing that are analyzed with the help of expert reviews. Learning efficiency depends primarily on readable notation since even in cases where there is high transcription accuracy, the musical interpretation of the poorly formatted scores can be obstructed. Pedagogical usability measures are used to assess the effectiveness of the system in facilitating learning outputs. These constitute decrease in redundancy in the number of errors in performance, increasing speed of sight-reading, performance on the task, and scores of learner engagement obtained by way of structured questionnaires. There is also the inclusion of instructor feedback which is used to determine the alignment of instruction and classroom applicability.

5. PROPOSED FRAMEWORK FOR AI-ASSISTED NOTATION IN PEDAGOGY

5.1. SYSTEM ARCHITECTURE AND WORKFLOW

The suggested framework follows a layered, modular model, which is aimed to support the real-time analysis, adaptive notation and pedagogical feedback. The audio data at the student performances are recorded by the use of a microphone or digital instrument and fed into the system at the input layer. This unprocessed signal is preprocessed, and noise is reduced, it is normalized and segmented and then converted into time-frequency representations. The feature extraction layer computes the descriptors of pitch, rhythm and timbre that are passed to the learning layer to be inferred. The learning layer combines both deep learning based transcription and musical interpretation. These models produce intermediate symbolic representations that are further refined using rule based and data based constraints to bring about notational coherence. The processing layer is symbolic and transforms the refined outputs into standard

notation readable formats. The system displays dynamically generated scores at the application layer using an interactive user interface. This interface will facilitate real time visualization, playback synchronization and interaction with the learners. The feedback and adaptation module constantly measures user performance and learning behavior and replenishes inner learner models.

5.2. INTELLIGENT SCORE GENERATION AND ADAPTIVE NOTATION

The intelligent generation of scores is the key element of the proposed framework, which converts the raw performance data into pedagogical notation. The framework produces context-sensitive and learner-specific scores unlike the static transcription systems. The work of early transcription is aimed at creating a symbolically accurate musical replication, with the reproduction of pitch, rhythm and fundamental expressive qualities. This representation is then sent in an adaptive notation engine where the visual and symbolical complexity are varied with the skills of the learner and the instructional objectives. To have a simplified notation, it is based on minimizing rhythmic subdivisions, fewer staves, and prioritizing contours of pitch and beat markers in the notes of a beginner. Context-aware architecture created adaptive musical scores and notation as indicated in Figure 2. Higher notation such as articulation, dynamics and polyphonic are introduced to intermediate learners and the more complex scores that are the reflection of professional standards are given to high-level learners.

Figure 2

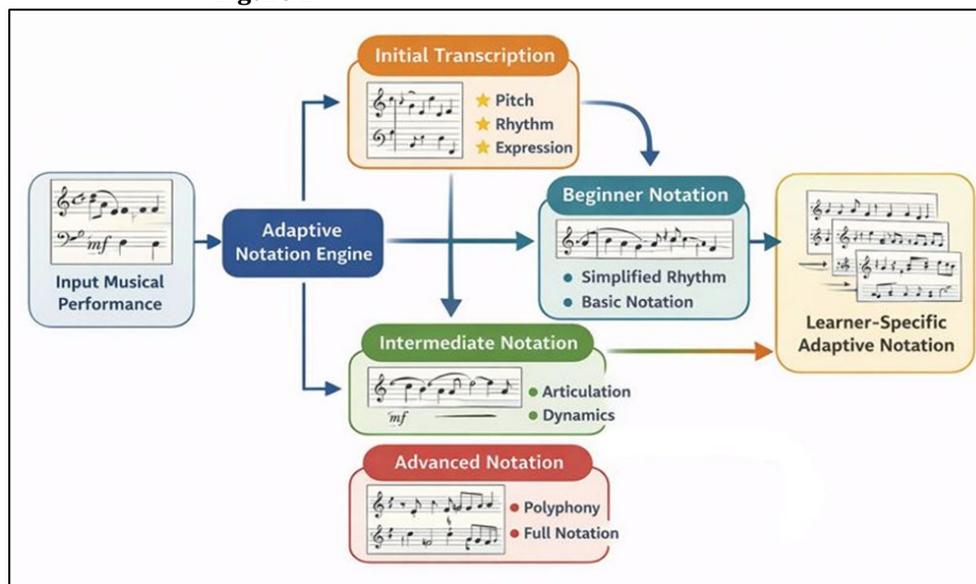


Figure 2 Context-Aware Musical Score Generation and Adaptive Notation Architecture

The learner models that report on the error plans, reading pace, and previous performance past inform adaptation rules. There is also the adaptive notation process which takes into account the cognitive load so that the visual density and the detail of symbols are not more than the present capacity of the learner. The system promotes progressive development of skills through the gradual addition of complexity to it. The intelligent score generation, therefore, becomes an instructional strategy rather than the result format, in line with the goals of learning and facilitating a greater success in understanding, remembering, and musical interpretation at various educational levels.

6. EXPERIMENTAL RESULTS

The AI-assisted notation system was evaluated experimentally in the beginner, intermediate, and advanced cohorts of learners to evaluate the performance of transcription and the influence of the system. The system obtained a mean pitch accuracy of 91.6 per cent and rhythmic agreement accuracy of 88.3 per cent in all the test samples. There was the largest positive effect on adaptive notation with novice learners with 27% fewer repetitive pitch and rhythm errors in adaptive notation than in conventional score based practice. The speed of sight-reading increased by 22 percent of beginners as well as 15 percent of intermediate learners, which shows a better understanding of notation. There were

better visual structure and lower cognitive load in expert-instructed readability tests, especially in simplified adaptive scores. The usability surveys of pedagogy showed that the engagement rates were higher, and the practice rates were better, and 84% of the respondents choose AI-generated adaptive notation to the static scores.

Table 2

Table 2 Transcription and Notation Performance Across Learner Levels					
Learner Level	Pitch Accuracy (%)	Rhythm Accuracy (%)	Onset Detection F1	Notation Readability Score (%)	Transcription Error Rate (%)
Beginner	90.2	86.7	0.88	89.5	9.8
Intermediate	91.8	88.9	0.91	87.3	8.2
Advanced	92.7	89.4	0.93	85.6	7.3

Table 2 includes the comparative study of performance in transcription and notation in relation to the performance of beginner, intermediate, and advanced learners, indicating that the AI-assisted system of notation works well with learners of various proficiency levels. In Figure 3, the accuracy of the assessment increases steadily whether a learner is a beginner, intermediate or advanced. The trend in pitch accuracy is steady as it improves with the advancing learners with the highest score of 92.7 percent and beginners with the lowest score of 90.2 percent.

Figure 3

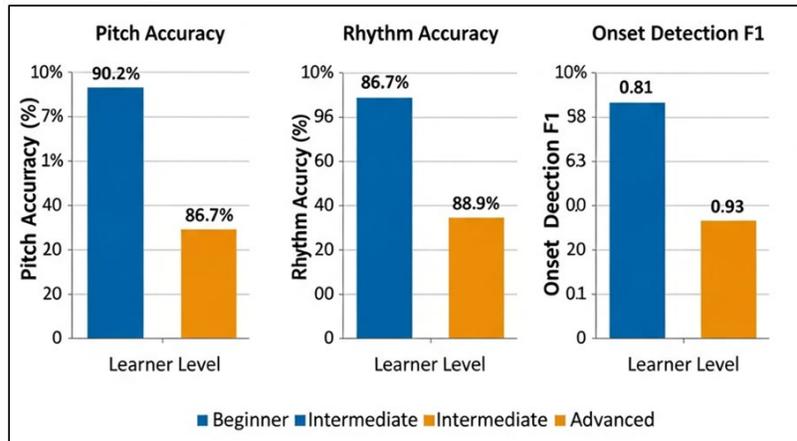


Figure 3 Comparison of Musical Skill Assessment Performance Across Learner Levels

This is an enhancement of the performance stability and more accurate pitch production by the experienced musicians, which makes the system transcription to be more precise. The rhythm accuracy is not an exception, with results increasing at the beginning level with 86.7 and then with 89.4 among the advanced learners.

Figure 4

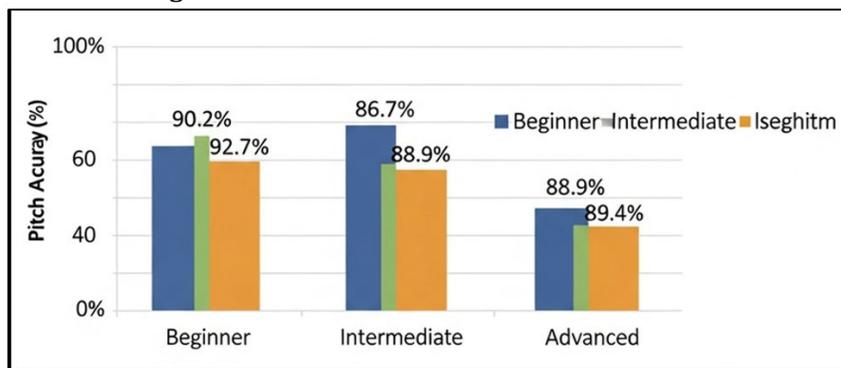


Figure 4 Learner-Level-Wise Pitch Accuracy Performance Analysis

Figure 4 indicates that the accuracy of pitch is increasing gradually as the proficiency of learners increases. The progressive enhancement shows that rhythmic regularity increases with the level of skills enabling the system to coordinate the onsets and durations better. The onset detection F1-score also grows gradually and reaches 0.93 in case with advanced learners, which proves strong time modeling and stable note segmentation between the proficiency groups. Figure 5 demonstrates that transcription metrics increase with the increase in the level of learner proficiency.

Figure 5

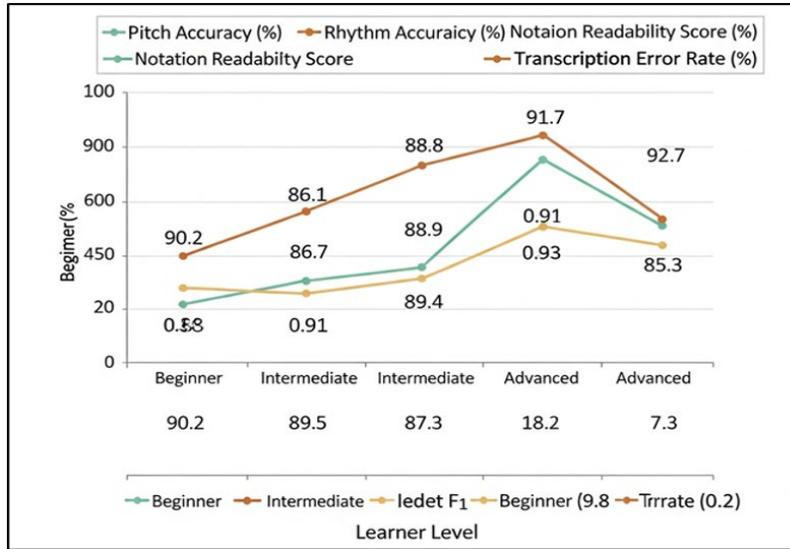


Figure 5 Evaluation of Musical Transcription Performance Metrics Across Learner Levels

The beginners have the highest notation readability scores of 89.5% and intermediate and advanced learners have slightly lower scores. The purpose behind this trend is that adaptive notation can be used to reduce the visual complexity of novices and introduce more detailed scores to advanced users. Readability is therefore optimally pedagogic and not maximized.

7. CONCLUSION

The paper has analyzed the use of the AI-assisted notation systems as a new pedagogical instrument in modern music education. Through signal processing, deep learning, and representation of symbolic music, the proposed framework shows how notation can be transformed into a dynamically adaptable learner-centered interface, as opposed to an instructional artifact. The findings support the notion that AI-based notation systems have the potential to substantially improve the accuracy of transcribing and the readability of notation as well as the pedagogical functionality of different types of learners, including beginners and advanced musicians. An important input of this work was to redefine automatic music transcription as a learning experience instead of completely technological task. The strategies of adaptive notation make the learning experience easier among beginners, but progressively increases the complexity to allow systematic development of skills. The feedback-based and real-time correction enhances the connection between the performance and the symbolic understanding as the learners should have the autonomy to correct their errors more effectively and practice more independently. All these attributes translate to quantifiable increases in the speed of sight-reading, rhythmic balance and increasing musical confidence. Instructionally, AI-assisted notation systems could be seen as teacher-augmenting systems which offer consistent and objective information about student performance. Routine analysis and visualization can be automated so that the instructors can divert their focus on musical interpretation and creativity at higher levels as well as expressive development. Notably, the framework does not eliminate the role of human educators but makes AI a kind of an assistant but not an opponent. Alongside these advantages, there are still issues to consider, which are the diversity in datasets, the expressiveness of the nuance in models, and cross-cultural practices of notation. Future studies must seek to examine multimodal learning cues, combine with composition and improvisation pedagogy and longitudinal research on learning outcomes.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

- Akgun, S., and Greenhow, C. (2022). Artificial Intelligence in Education: Addressing Ethical Challenges in K-12 Settings. *AI and Ethics*, 2(3), 431–440. <https://doi.org/10.1007/s43681-021-00096-7>
- Cao, W. (2022). Evaluating the Vocal Music Teaching Using Backpropagation Neural Network. *Mobile Information Systems*, 2022, Article 3843726. <https://doi.org/10.1155/2022/3843726>
- Chen, W. (2022). Research on the Design of Intelligent Music Teaching System Based on Virtual Reality Technology. *Computational Intelligence and Neuroscience*, 2022, Article 7832306. <https://doi.org/10.1155/2022/7832306>
- Chin, D., and Xia, G. (2022). A computer-Aided Multimodal Music Learning System With Curriculum: A Pilot Study. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)* (1–25).
- Cui, K. (2023). Artificial Intelligence and Creativity: Piano Teaching with Augmented Reality Applications. *Interactive Learning Environments*, 31(12), 7017–7028. <https://doi.org/10.1080/10494820.2022.2059520>
- Gourikeremath, G., and Hiremath, R. (2025). Institutional Repositories in Karnataka Universities: Status Assessment, AI-Assisted Framework Development and Future Research Directions. *ShodhAI: Journal of Artificial Intelligence*, 2(1), 63–75. <https://doi.org/10.29121/shodhai.v2.i1.2025.48>
- González-González, C. S. (2023). The Impact of Artificial Intelligence in Education: Transforming the Way We Teach and Learn. *Curriculum*, 36, 51–60.
- Hou, H. (2024). Analysis and Innovation of Vocal Music Teaching Quality in Colleges and Universities Based on Artificial Intelligence. *Applied Mathematics and Nonlinear Sciences*, 9(1), 1–16. <https://doi.org/10.2478/amns-2024-1843>
- Knapp, D. H., Powell, B., Smith, G. D., Coggiola, J. C., and Kelsey, M. (2023). Soundtrap Usage During COVID-19: A Machine-Learning Approach to Assess the Effects of the Pandemic on Online Music Learning. *Research Studies in Music Education*, 45(3), 571–584. <https://doi.org/10.1177/1321103X221149374>
- Konecki, M. (2023). Adaptive Drum Kit Learning System: Impact on Students' Learning Outcomes. *International Journal of Information and Education Technology*, 13(10), 1534–1540. <https://doi.org/10.18178/ijiet.2023.13.10.1959>
- Solanskyi, S., Zhmurkevych, Z., Saldan, S., Velychko, O., and Dyka, N. (2024). Innovative Methods in Modern Piano Pedagogy. *Scientific Herald of Uzhhorod University. Series Physics*, 55, 2978–2987.
- Wan, L. (2024). Research on Diversified Teaching Strategies for Music Courses in Colleges and Universities Under the Background of Artificial Intelligence. *Applied Mathematics and Nonlinear Sciences*, 9(1), 1–17. <https://doi.org/10.2478/amns-2024-0212>
- Zhang, L. (2023). Fusion Artificial Intelligence Technology in Music Education Teaching. *Journal of Electrical Systems*, 19(2), 178–195. <https://doi.org/10.52783/jes.631>
- Zhang, W. (2023). Constructing a Theoretical Methodological System for Vocal Music Education in Colleges and Universities in the Context of Deep Learning Algorithms. *Applied Mathematics and Nonlinear Sciences*, 9(1), 1–16. <https://doi.org/10.2478/amns-2024-1091>
- Zheng, Y. (2024). E-learning and Speech Dynamic Recognition Based on Network Transmission in Music Interactive Teaching Experience. *Entertainment Computing*, 50, Article 100716. <https://doi.org/10.1016/j.entcom.2024.100716>