Original Article
ISSN (Online): 2582-7472

ShodhKosh: Journal of Visual and Performing Arts
Special Issue on Digital Preservation and Intelligent Innovation in Traditional and Modern Arts 7(1s)

MANAGING CROSS-CULTURAL MUSIC LEARNING WITH AI TOOLS

Dr. Shweta Bajaj [1], Om Prakash [2], Amol Barde [3], Damodaran B. [4], Chaitali [5], Monali Gulhane [6]

[1] Associate Professor, School of Management and School of Advertising, PR and Events, AAFT University, Raipur, Chhattisgarh-492001, India
[2] Associate Professor, School of Business Management, Noida International University, Greater Noida 203201, India
[3] Department of Mechanical Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, 411037, India
[4] Associate Professor, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu 600094, India
[5] Department of Computer Science and Engineering, Shri Shankaracharya Institute of Professional Management and Technology, Raipur, Chhattisgarh, India
[6] Symbiosis Institute of Technology, Nagpur Campus, Symbiosis International (Deemed University), Pune, India

Received 13 September 2025
Accepted 12 December 2025
Published 17 February 2026

Corresponding Author
Dr. Shweta Bajaj,
shweta.bajaj@aaft.edu.in

DOI
10.29121/shodhkosh.v7.i1s.2026.7112

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

ABSTRACT

Teaching cross-cultural music is a pedagogical problem of cross-cultural differences in musical structure or notation, ornamentation, rhythmic patterns, and linguistic articulation. New opportunities to overcome these challenges are introduced by recent advances in artificial intelligence (AI) that allow making music education settings adaptive and culturally informed. This paper gives a proposal of an AI-based framework to regulate the process of learning cross-cultural music by introducing music information retrieval (MIR) models, Transformer-based architecture, generative models, and intelligent feedback mechanisms. The framework is aimed at the learners and educators who work with different musical traditions like the Western, non-Western, indigenous and folk music systems. The suggested system includes the recognition of style and ornamentation to locate the culturally specific melodic and rhythmic patterns, cross-cultural notation and rhythm transformation to overcome the differences in the representational systems, and AI-provided feedback on the pronunciation, phrasing and expressiveness nuances. To make sure that it is pedagogically valid and culturally sensitive, several groups of users are involved including students, teachers and experts in cross-cultural music. Measurements against technical accuracy are not limited to evaluate the achievement, but cultural fidelity, interpretive correctness, and quantifiable learning outcomes are considered the evaluation metrics. The system has proven to be relevant in East-West classical fusion classes, AI-driven orchestration of indigenous and folk repertoires and cross-language syllabus voice training and lyric-alignment.

Keywords: Cross-Cultural Music Education, Artificial Intelligence in Music, Music Information Retrieval, Cultural Fidelity, AI-Assisted Learning Systems, Global Music Pedagogy

1. INTRODUCTION

Music is a universal form of human expression, which is, nevertheless, strongly influenced by the cultural setting, historical perspectives, linguistic attributes and social definitions. Every musical culture has its own unique musical

systems of melodies, rhythm, tuning, decoration, notation and performance practice. With the growing globalization of music education, students and teachers are progressively experiencing interaction with other musical cultures other than their own. The cross-cultural interaction is very creative, and educational and at the same time poses serious challenges in terms of pedagogy. Students usually have difficulties in receiving the unknown rhythmic cycles, microtonal intervals, stylistic ornaments, pronunciation patterns, and expressive conventions which cannot be easily communicated with the help of a standard method of teaching and standard notation systems. The paradigm of traditional music pedagogy has mainly been based on human teachers, oral learning and apprenticeship system based on the culture to pass on these subtle aspects. Although they work exceptionally well in certain traditions, these methods become ineffective in cross-cultural situations, especially in formal education systems where time and expertise in addition to access to all varieties of cultural experts are limited. In addition, written notation and non-specific digital learning environments tend to favor the Western musical standards, and therefore discriminate against non-Western, indigenous, and folk musical traditions, inadvertently Li and Wang (2024). These limitations point to the necessity of having new learning devices that can enable culturally adaptive and sensitive as well as scalable learning of music in different traditions. Artificial intelligence (AI) holds potential solutions to these problems due to recent developments. Artificial intelligence technologies, especially, music information retrieval (MIR), deep learning, Transformer-based sequence models, and generative systems have shown great potential in performing analysis of musical structure, modeling performance patterns, and real-time feedback. When used in an intelligent way, these instruments have the potential to go beyond surface-level precision to represent culturally-specific musical characteristics of ornamentation styles, rhythmic expression, tonal inflections, and expressive timing Civit et al. (2022).

This opens the potential of AI systems that do not simply provide lessons on the notes, but assist the learners in the process of understanding how music operates in the context of its culture and expressiveness. The AI can be used as a smart broker of traditions in the field of cross-cultural learning of music. Through the training on a wide range of datasets of different musical cultures, AI systems have the right to aid automated style recognition, intercultural notation and rhythm changeover, and adaptive responses to definite traditions Agostinelli et al. (2023). As an example, an art student who has studied Western classical music may engage in the exploration of raga-based improvisation, local rhythmic patterns, or folk vocal ornamentation with the aid of AI tools and get feedback without dictating to him or her what is right or wrong. Likewise, teachers will also be able to use AI to add value to teaching, visualize intricate aspects of the culture, and create inclusive curriculum that is representative of global music. But, there are also important questions with the implementation of AI in the cross-cultural music education. Cultural representation, authenticity, and interpretive accuracy should be handled with a lot of care to prevent oversimplification, bias or cultural appropriation. AI systems should hence be modeled in a clear consideration of cultural knowledge representation, expert validation and pedagogical intent Copet et al. (2023). Such metrics as cultural fidelity and interpretive correctness are valued as much as the technical measurements of performance such as pitch or rhythm accuracy.

2. LITERATURE REVIEW

2.1. CROSS-CULTURAL PEDAGOGY IN MUSIC EDUCATION

Music education Cross-cultural pedagogy focuses on learning where the different musical traditions, values, and knowledge transmission ways are recognized, respected and incorporated. Initial contributions to the field were through ethnomusicology and multicultural education, which argued that non-Western, indigenous, and folk forms of music should be included in Western classical-small, classical curriculum. The main idea of cross-cultural pedagogy is that musical meaning cannot be discussed outside of its cultural, social, and linguistic context Huang et al. (2023). Consequently, to teach appropriately, it is not enough that one should be educated technically, but it is necessary to appeal to stylistic and oral traditions, performance, and cultural narratives. Scholars point out the difficulties when using cross-cultural music education, especially in formal institutions. Educators might not be experts in several of the traditions, and the standardized notation and evaluation systems have a tendency to favor the western models. Oral communication, improvisation and learning through the community, which is practiced in most societies, find it hard to comply with the traditional classroom model Schneider et al. (2023). It is also observed that learners may have problems in adjusting to new tonal systems, rhythmic patterns, and expressive rules and norms that could make them understand the information superficially instead of meaningfully. Experiential learning, intercultural dialogue, and reflective practice are some of the strategies promoted by the contemporary pedagogical models.

2.2. AI TOOLS FOR MUSIC ANALYSIS, NOTATION, AND PERFORMANCE FEEDBACK

The application of artificial intelligence to music analysis, notation and performance feedback has become more frequent as a result of innovations in music information retrieval (MIR), deep learning and sequence modeling. The first AI systems, in fact the first, were task-based, including pitch recognition, time determination, and symbolic representation, mainly of Western tonal music. More recent advancements though have extended these features to allow rhythmic complexity analysis, expressive timing, ornamental detection and timbral characterising Ning et al. (2025). One application that has been specifically successful is transformer-based network and recurrent neural networks to model long-range musical dependencies and performance variations. The AI-assisted notation tools can be used to provide automated transcription, alignment of scores, and adaptive visualization of the musical structure; the tools lower the learning barrier of students with learning difficulties in traditional notation. The systems of performance feedback use audio analysis to give real-time or post-performance feedback regarding pitch accuracy, rhythm stability, articulation, and dynamics Yu et al. (2024). These systems have recorded good effects on the engagement and efficiency of learners, particularly in the self-directed learning settings.

2.3. CULTURAL KNOWLEDGE REPRESENTATION IN AI SYSTEMS

The representation of knowledge on cultural aspects in AI systems deals with the question of the possibility to encode symbolic meaning, contextual knowledge, and rules that are specific to a certain tradition in computational patterns. This, in the music-related AI, is represented not only in the acoustic terms but also in the culturally-based ideas like the stylistic grammar, performance traditions, linguistic relations, and social role. The initial methods used were based on rule-based models and ontologies, which were manually defined to represent the cultural structures. Although interpretable, they were not very scalable and adaptable between traditions Marquez-Garcia et al. (2022). Modern AI studies are starting to use more data-driven techniques, in which cultural trends are implicitly learned based on large, annotated datasets. In music, this involves acquisition of stylistic signatures both through records, score, lyrics and performance meta data. Hybrid methods that use symbolic representations together with neural models are beginning to be considered, with the advantage that they are less interpretable and more expressive. According to Table 1, AI can facilitate the efficient learning of cross-cultural music with the help of adaptive technologies. With these techniques, AI can be used to match low-level signal characteristics to high-level cultural representations, e.g. mode, type of ornamentation, ritual situation. Nevertheless, ethical and methodological issues are also mentioned in the literature. The knowledge of cultures is usually informal, shifting and contentious and can hardly be formalized without simplifying it Eftychios et al. (2021).

Table 1

Table 1 Summary on AI-Supported Cross-Cultural Music Learning

Musical Domain	AI Technique Used	Cultural Scope	Learning Task	Limitations
Western Classical	MIR + HMM	Western	Pitch & rhythm training	No cultural adaptation
Chinese Classical	CNN	Single-culture (East)	Ornament recognition	Limited learner feedback
Folk Music	Rule-based + MIR	Regional folk	Pattern annotation	Not learner-oriented
Vocal Music Ruth and Müllensiefen (2021)	RNN	Mono-lingual	Singing pronunciation	Language-specific only
Indian Classical	Deep CNN	Single-culture (South Asia)	Raga identification	No pedagogy integration
Music Education	Transformer	Western-centric	Score following	Western notation bias
Music Performance	MIR + ML	Multi-genre	Expressive timing	Lacks cultural semantics
Vocal Training Wei et al. (2022)	AI feedback system	Cross-language	Lyric alignment	Cultural expression ignored
Music Pedagogy	Hybrid AI	Western & Jazz	Practice optimization	Limited cultural diversity
Indigenous Music Wu (2025)	AI annotation	Indigenous	Oral tradition support	Small datasets

Cross-Cultural Music	Transformer + MIR	East–West	Style transfer	No learning assessment
Multilingual Vocal	Speech–Music AI	Multi-language	Pronunciation learning	Weak expressive feedback
Multi-Tradition Li (2025)	MIR + Transformers + Generative AI	Global	Learning & interpretation	Requires large expert-validated datasets

3. METHODOLOGY

3.1. AI TOOLS USED: MIR MODELS, TRANSFORMERS, GENERATIVE MODELS, FEEDBACK SYSTEMS

The suggested approach uses a combination of several AI applications to consider analytical, generative, and pedagogical needs of learning music across cultures. MIR models constitute the bottom layer, which allows retrieving low-level and mid-level musical features, e.g., pitch contours, rhythmic patterns, timbral features, and ornamentation patterns. They are trained on culturally diverse data to learn to embrace variations between the musical traditions of different cultures, such as non-Western tonality systems and more complicated rhythm patterns. The feature representations that are produced by MIR modules are the inputs to the higher-level learning components Alexander (2024). Long-range temporal dependencies in melodic and rhythmic sequences are modeled with the help of transformer-based architecture. The system enables them to capture contextual relations that are important in their understanding of phrasing, improvisation, and stylistic continuity across cultures because of their self-attention mechanisms. Synthesizing culturally informed examples of music, other ornamentations, and transformed notations are generated via generative models, such as sequence-to-sequence and diffusion-based models, and can aid exploratory and creative learning. Figure 1 demonstrates unified MIR, transformer, generative models, which allow one to learn music across cultures. Through these models, learners can juxtapose the stylistic interpretations of the traditions and at the same time, the structural coherence is sustained. The pedagogical interface of the framework is made up of feedback systems. The system uses adaptive feedback of the use of pitch, interpretation of rhythm, pronunciation, and expressive nuances by converting real-time audio data analysis with cultural representations learned.

Figure 1

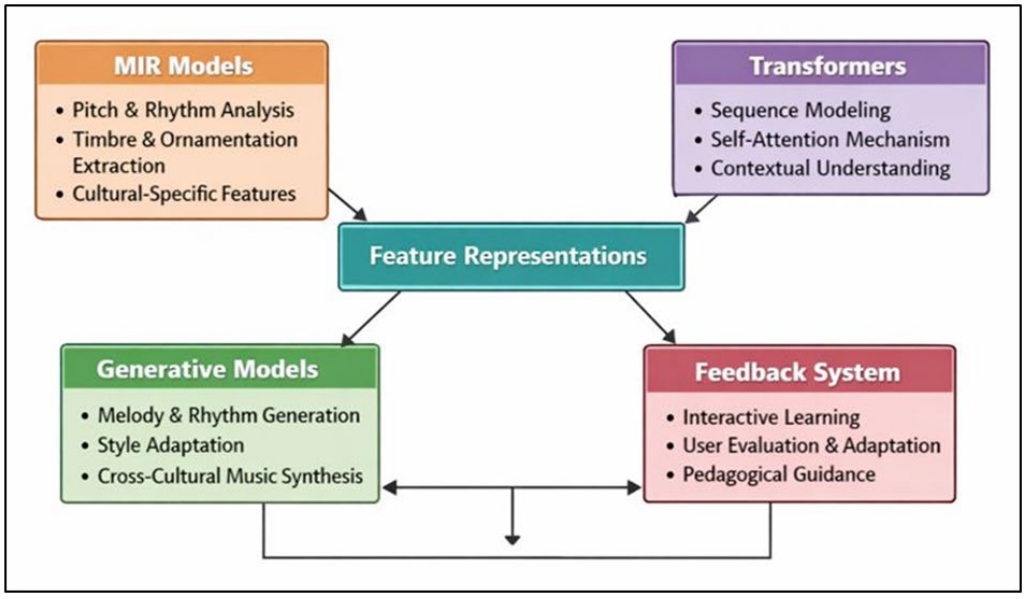

Figure 1 AI Framework Integrating MIR, Transformer, and Generative Models for Cross-Cultural Music Learning

Significantly, feedback levels and appraisal reasoning is also adapted to culture, without imposing the strict Western performance standards. These AI tools make up a combined learning environment that helps in analyzing, generating, and culturally aware guidance in the cross-cultural learning of music.

3.2. USER GROUPS: STUDENTS, TEACHERS, AND CROSS-CULTURAL EXPERTS

To make sure that the pedagogical design is relevant, useful, and culturally authentic, the methodological design precisely involves three major users of the design: students, teachers, and experts on cross-cultural matters. Students are the representatives of different musical abilities and different degrees of musical backgrounds, including beginners who should meet with the new traditions and high-level learners who wish to specialize in cross-cultural direction Gruetzemacher and Whittlestone (2022). They are engaged in the system in terms of guided practice, exploratory learning, and self-assessment. User interfaces are structured to support the different literacies in notation, language, and technology, to allow all individuals to contribute towards cultural backgrounds. As facilitators and Desai (2024) mediators of learning, teachers are in the center stage. The AI system is not supposed to substitute teaching power, but it is rather introduced as a supportive tool and assists in lesson planning, demonstration, and formative assessment. In addition to adapting learning modules, choosing cultural reference models, and decoding AI-generated feedback to meet curricular objectives Lee (2024), teachers are able to do so. Such design gives a recognition to the fact that pedagogical judgment and contextual explanation are still required especially in case of culturally sensitive material. Ethnomusicologists, tradition bearers and linguistic specialists are cross-cultural experts whose knowledge of domains will be helpful in system training and validation.

3.3. EVALUATION METRICS: CULTURAL FIDELITY, LEARNING OUTCOMES, INTERPRETIVE ACCURACY

The assessment in the suggested methodology goes beyond the usual technical indicators to respond to the cultural and pedagogical objectives of cross-cultural music learning. Cultural fidelity is used to assess how AI-generated analyses, feedback and musical work are in line with the stylistic conventions, expressive rules and contextual significance of a particular tradition. This measure is evaluated by expert judgment, comparison tests of listening, and agreement between AI interpretation of the work and human cultural judgments. It is focused on the respect of stylistic flexibility of most traditions, as opposed to strict correctness. Quantitative and qualitative indicators are used to test the learning outcomes. Measures in the quantitative category are such as the improvement in pitch stability, rhythmic accuracy, consistency in pronunciation, and effectiveness of practice with time. The qualitative tests demonstrate the learner confidence, cultural awareness, and reflective understanding in the form of the survey, interviews, and performance reviews. These results give a clue about the ability of the system to favor acquisition of skills and intercultural musical competence. Interpretive accuracy puts emphasis on the skill of the learner to play and explain culturally suitable musical interpretations.

4. AI-ENABLED CROSS-CULTURAL MUSIC LEARNING SYSTEM

4.1. AUTOMATED STYLE AND ORNAMENTATION RECOGNITION

The proposed AI-based cross-cultural music learning system will rely on automated style and ornament recognition as one of its fundamental elements, covering a highly problematic area of intercultural music education. The ornamentation, like last notes of grace, slides, variations of vibratos, microtones, or rhythmic embellishments bear a great stylistic and cultural sense. These aspects are passed down orally and are hard to notationate as well as taught by rule. The system uses the deep learning models and music information retrieval to detect and label ornamentation patterns using only audio recordings and symbolic notations of ornamentation. Through the processing of pitch curves, time variation, and spectral features, the AI models can acquire unique stylistic features in relation to a particular musical tradition and school of music performance. Modeling of sequences with transformers makes it possible to model long-range dependencies so that ornamentation can be understood within melodic context and phrasing and not as an isolated event. Such contextual awareness is imperative to the differentiation of the expressive variation and structural features. Automated recognition serves as an analytical and pedagogical tool to the learners. Auditory and visual feedback points out the violations of style, providing examples of culturally appropriate ornamentation and the performance of the learner.

4.2. CROSS-CULTURAL NOTATION AND RHYTHM TRANSFORMATION

The learners are able to travel across divergent systems of musical representation systems even with cross-cultural notation and rhythm transformation keeping cultural meaning intact. There are many musical traditions, with different conceptualizations and notations of rhythm, pitch and structure, and including staff-based Western notation, mnemonic syllables, tablatures, and oral rhythmic cycles. The offered system uses AI-enhanced mapping to convert musical content between these systems, which lowers the cognitive load of the learners who are trained in one of them. In rhythm, AI models detect internal rhythmic patterns and cycles so that one can convert complex rhythmic systems without changing accentuation and phrasing. In the case of pitch, microtonal and modal variations are made possible by the system since the system does not make any strict assumptions based on equal temperament. This teaching is important in terms of pedagogy. Students have a chance to see the same musical text in various representational perspectives, which allows them to provide comparative analysis and intercultural understanding. Educators have the opportunity to use representations that will suit the background of learners and slowly suggest different notational views.

4.3. AI-GUIDED PRONUNCIATION, PHRASING, AND EXPRESSIVE ELEMENTS

Pronunciation, phrasing, and expressive aspects are guided by AI and tackle the vocal and expressive aspects of cross-cultural music learning, where the expressions and language elements are closely intertwined. Vocal traditions have tended to combine phonetics, accentuation and syllabic timing that are language specific with musical structure, thus correct pronunciation and phrasing is a key to sound performance. The suggested system combines models of speech and singing analysis along with a group of music-specific AI elements in order to deliver culturally informed vocal feedback. Pronunciation analysis aims at phoneme accuracy, vowel shaping and consonant articulation in the conditions of singing considering the differences across languages and singing styles. The system also matches melodic contours and rhythmic placement with the lyrics allowing timing and stress pattern to be accurately viewed. Modeling phrasing and expression is based on the study of dynamics, flexibility of timing and shaping of melodic contour, expressive norms specific to a given culture, e.g., rubato, call-and-response, or narrative. Instead of using standard measures of voices, the AI tends to tone its response to the cultural system chosen. Interpretive suggestions are presented to learners based on stylistic examples which may be alternative phrases forms or use of timing.

5. CASE STUDIES AND IMPLEMENTATION

5.1. EAST–WEST CLASSICAL FUSION LEARNING MODULES

As the case study of East West classical fusion shows, the proposed AI-enabled system will empower the learners to work with amalgamated musical forms that will combine western classical modes with non-western melodical and rhythmic traditions. Fusion learning modules are structured on the principles of parallel presentation of the stylistic aspects, which gives the learners an opportunity to discover the relationships between harmonic progression, modal structures, ornamental features, and rhythmic phrasing. The AI-based style recognition can determine culturally specific aspects in each tradition and illustrate the points of agreement and disagreement in the fusion pieces. Figure 2 indicates that AI is capable of facilitating systematic learning on the fusion of East and western classical music. Students will work with modular activities that processively incorporate both systems in them.

Figure 2

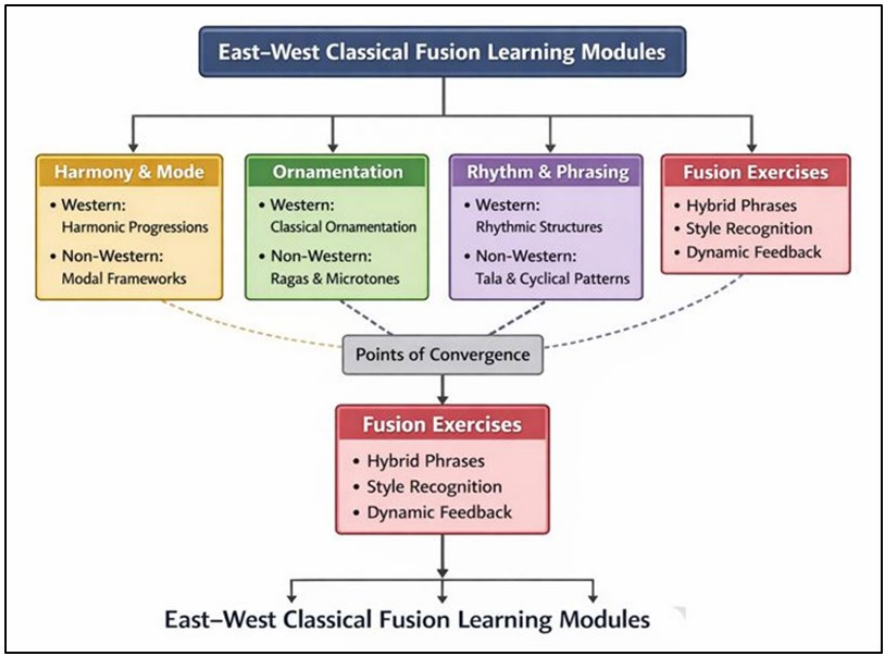

Figure 2 AI-Enabled East–West Classical Fusion Learning Framework

As the case study of East West classical fusion shows, the proposed AI-enabled system will empower the learners to work with amalgamated musical forms that will combine western classical modes with non-western melodical and rhythmic traditions. Fusion learning modules are structured on the principles of parallel presentation of the stylistic aspects, which gives the learners an opportunity to discover the relationships between harmonic progression, modal structures, ornamental features, and rhythmic phrasing. The AI-based style recognition can determine culturally specific aspects in each tradition and illustrate the points of agreement and disagreement in the fusion pieces. Figure 2 indicates that AI is capable of facilitating systematic learning on the fusion of East and western classical music. Students will work with modular activities that processively incorporate both systems in them.

5.2. INDIGENOUS AND FOLK MUSIC LEARNING SUPPORTED BY AI ANNOTATION

The second case study is indigenous and folk music learning, in which the oral tradition, community, and variability of style are the main aspects. The AI annotation tools are used to complement the documentation, learning, and interpretation rather than disrupting the conventional pedagogical methods. Audios are marked cultural significances, such as phrase marks, rhythmic patterns, ornamentation styles and performance environment and contextual footnotes. To help the learners, the AI system allows annotating these passages, then playing them up and down and practicing something specifically. Instead of fixed scores, the system is focused on flexible representations, representing oral traditions. Learners are able to change the speed of the recording, isolate stylistic elements, and compare the performance of the same piece in several different ways to learn to tolerate certain variation. Pattern recognition with the usage of AI emphasizes repetitive features and still allows seeing unique expressive variations. The community professionals and bearers of tradition will be involved in the process of designing annotations and verifying them, making sure that cultural meanings are properly reflected. Educators combine AI markups with blended learning, which is a combination of classroom teaching and self-directed learning. The results of the implementation include increased engagement of the learners, better retention of the stylistic elements, and became more respectful of the cultural context. This case study shows that AI annotation can be an auxiliary catalyst between the old system of knowledge and the new settings of schools and colleges.

5.3. CROSS-LANGUAGE VOCAL TRAINING AND LYRIC ALIGNMENT

The third case study focuses on cross-language vocal training, which deals with issues of learners singing in foreign languages. Vowel authenticity, syllabic timing and emotive expression play a vital role in the vocal quality but not easily acquired without exposure to native languages. The AI system incorporates the alignment of the lyrics, phonetics, and melodic mapping to assist in culturally aware learning of the voice. The melodic and rhythmic structures are written at the syllable and phoneme level and allow accurate feedback regarding cursory and Stress patterns. That system recognizes the deviations of pronunciation and offers corrective suggestions based on the phonetics of singing and not the norms of the spoken language. Students are presented with visual indications and relative audio samples, which makes them minimize progressively without the disheartening effect of expressive experimentation. The system allows teachers to create multilingual vocal tasks and track the development of learners in the areas of pronunciation accuracy and expressive delivery. Linguistic models are tested by cultural experts in order to provide proper articulation and style refinement. The outcome of the implementation indicates that there is a great improvement in the sound clarity, rhythmic consistency and confidence. Students also add that they are more culturally aware and less anxious about performing as new speakers. In this case study, we can emphasize on the fact that the system can combine linguistic and musical aspects, and it can facilitate the inclusion and authenticity of cross-cultural vocal learning.

6. RESULTS AND ANALYSIS

The results of the experimental assessment of three case studies prove the consistent improvement of the technical performance, cultural awareness, and engagement of a learner. There were observable positive changes in the stability of their pitch, rhythmic adjustment, and accuracy in pronunciation among learners with the system equipped with AI in contrast to basic instruction. What is more important, professional judgments were higher concerning fidelity to culture, especially the use of ornamentation, phrasing and stylistic balance in the situation of fusion. Differentiation of stylistic motifs and sensitivity to acceptable variation were found to be better with indigenous and folk modules. Vocal training in a different language led to decreasing the timing mistakes and phonetic inaccuracies and enhancing the expressiveness. Teachers said their instructional load was low and formative assessment enriched.

Table 2

Table 2 Learning Performance Comparison – Traditional vs. AI-Enabled Cross-Cultural Music Learning

Metric	Traditional Instruction	AI-Enabled Learning System	Improvement (%)
Pitch Accuracy (%)	78.9	92.6	13.7
Rhythm Alignment Score (%)	74.3	90.8	16.5
Ornamentation Accuracy (%)	61.5	85.2	23.7
Pronunciation Accuracy (%)	69.8	88.9	19.1
Practice Time to Master Task (min ↓)	46.2	31.4	–14.8

As shown in Table 2, cross-cultural music learning system based on AI has an evident and uniform benefit over the conventional instructional approach in all the measured aspects. The percentage of pitch accuracy increases by 78.9 to 92.6 which means that AI-based feedback and adaptive practice can lead to a significant improvement in tonal accuracy even in the situations when learners are required to work with novel musical regimes. Figure 3 indicates that AI-enabled learning is beneficial compared to traditional methods on important music education measures.

Figure 3

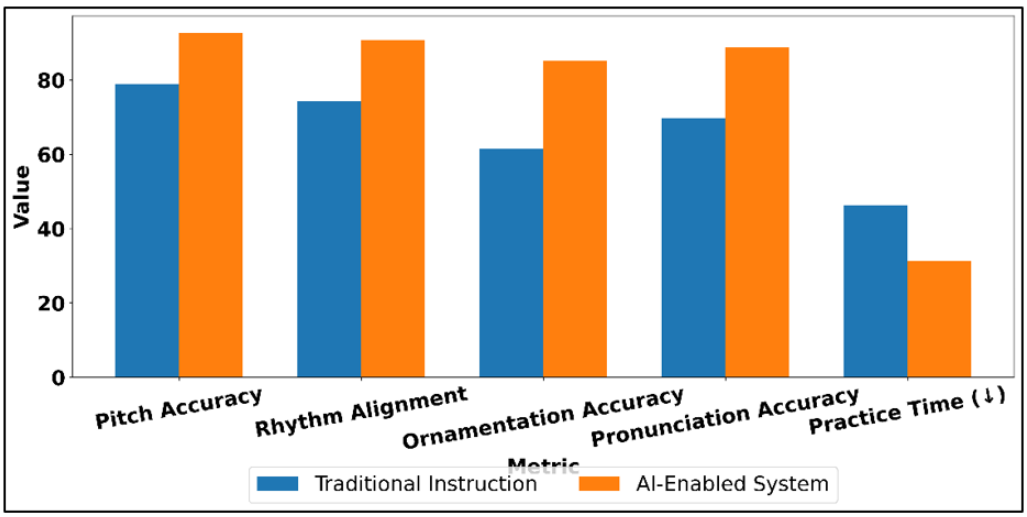

Figure 3 Comparison of Traditional vs AI-Enabled Music Learning Metrics

The same trend is witnessed in terms of rhythm alignment; the 16.5% increase is an indication that the system has the power to model and strengthen complicated rhythmic patterns that would otherwise be hard to learn using standard teaching methods. Among the most significant changes is in the ornamentality accuracy which increases by 23.7. Here the efficacy of automated style recognition and culturally sensitive feedback is pointed out as the means of transmitting subtle, tradition-sensitive additions that are generally passed on orally. There are also significant improvements in pronunciation accuracy, which is also in support of the importance of cross-language vocal training using combined lyric alignment and phonetic guidance. In Figure 4, the distribution of improvements in response to key music learning performance metrics is proportional. There is also the decrease in practice time necessary to learn a task, 46.2 to 31.4 minutes, which is an indication that learning was more efficient. Students have the advantage of specific feedback which is instant and minimal in terms of repetitive trial and error.

Figure 4

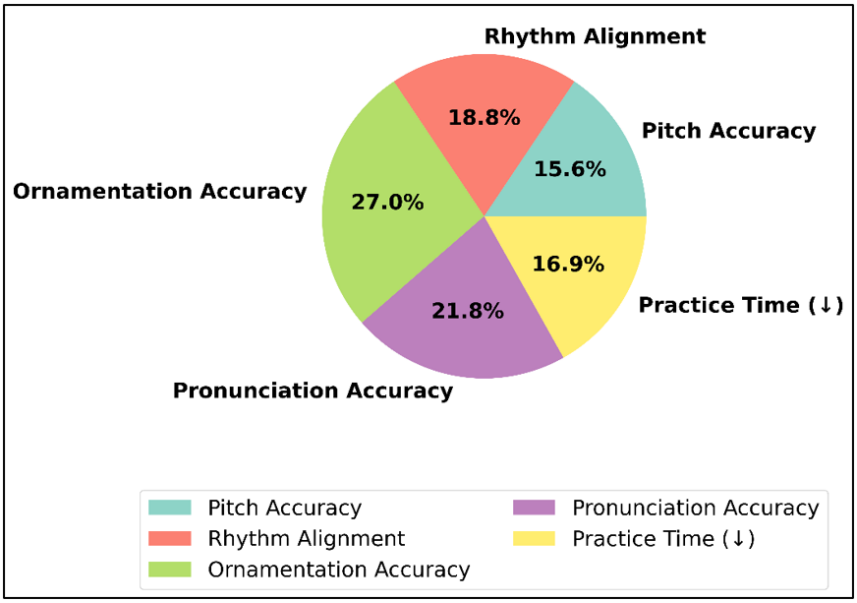

Figure 4 Pie Distribution of Improvement Across Music Learning Metrics

In general, these results imply that AI-based learning enhances technical and interpretative performance, as well as efficient use of practice, thus, increasing the accessibility, effectiveness and efficiency of cross-cultural music education.

7. CONCLUSION

This paper shows that artificial intelligence can be transformative in controlling learning of cross-cultural music in case it is designed with the purpose of pedagogy and cultural sensitivity. The proposed framework can be used to help learners navigate through various musical traditions, such as with the help of MIR models, Transformer architectures, generative systems, and adaptive feedback, without giving preference to a specific cultural norm. The findings reveal that AI may contribute to the improvement of technical precision, as well as, to interpretative richness, stylistic sensitivity, and expressive competence. Notably, the study places AI as a facilitating and complementary device, and does not supplant human knowledge. Pedagogical power lies in the hands of teachers, and the cross-cultural experts make sure that everything is authentic and ethical. Individualized instructions, comparative visualization, and culturally informed feedback are beneficial to the learners, which would be otherwise hard to provide at scale. The case studies demonstrate that AI may be effectively used to facilitate fusion learning, preservation-based folk education, and cross-language vocal training in the same system. Design considerations are also critical as found out. Cultural fidelity has to be referred to as a fundamental assessment criterion, along with learning outcomes and interpretive accuracy. A variety of datasets, professional participation, and adjustable levels of feedback is needed to prevent bias or simplification associated with the cultural aspect of the task. Obedience to these principles means that AI systems will be able to promote more profound cultural insight and not a shallow simulation.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

Agostinelli, A., Denk, T. I., Borsos, Z., Engel, J., Verzetti, M., Caillon, A., Huang, Q., Jansen, A., Roberts, A., Tagliasacchi, M., et al. (2023). MusicLM: Generating Music From Text (arXiv:2301.11325). arXiv.

Alexander, A. (2024). "Heart On My Sleeve": An AI-Created hit Song Mimicking Drake and the Weeknd Goes Viral (SAGE Business Cases Originals). SAGE Publications. https://doi.org/10.4135/9781071928646

Civit, M., Civit-Masot, J., Cuadrado, F., and Escalona, M. J. (2022). A Systematic Review of Artificial Intelligence-Based Music Generation: Scope, Applications, and Future Trends. Expert Systems with Applications, 209, 118190. https://doi.org/10.1016/j.eswa.2022.118190

Copet, J., Kreuk, F., Gat, I., Remez, T., Kant, D., Synnaeve, G., Adi, Y., and Défossez, A. (2023). Simple and Controllable Music Generation. Advances In Neural Information Processing Systems, 36, 47704–47720.

Desai, A. U. (2024). A Review of the Applications of Machine Learning in Cybersecurity and its Challenges. Journal of Digital Security and Forensics, 1(1), 26–29. https://doi.org/10.29121/digisecforensics.v1.i1.2024.17

Eftychios, A., Nektarios, S., and Nikoleta, G. (2021). Alzheimer Disease and Music-Therapy: An Interesting Therapeutic Challenge and Proposal. Advances in Alzheimer's Disease, 10, 1–18. https://doi.org/10.4236/aad.2021.101001

Gruetzemacher, R., and Whittlestone, J. (2022). The Transformative Potential of Artificial Intelligence. Futures, 135, 102884. https://doi.org/10.1016/j.futures.2021.102884

Huang, Q., Park, D. S., Wang, T., Denk, T. I., Ly, A., Chen, N., Zhang, Z., Zhang, Z., Yu, J., Frank, C., et al. (2023). Noise2Music: Text-Conditioned Music Generation With Diffusion Models (arXiv:2302.03917). arXiv.

Lee, E. (2024). Prompting Progress: Authorship in the Age of AI. Florida Law Review, 76, 1445. https://doi.org/10.2139/ssrn.4609687

Li, P.-P., and Wang, B. (2024). Artificial Intelligence in Music Education. International Journal of Human–Computer Interaction, 40, 4183–4192. https://doi.org/10.1080/10447318.2023.2209984

Li, S. (2025). The Impact of AI-Driven Music Production Software on the Economics of the Music Industry. Information Development. Advance online publication. https://doi.org/10.1177/02666669241312170

Marquez-Garcia, A. V., Magnuson, J., Morris, J., Iarocci, G., Doesburg, S., and Moreno, S. (2022). Music Therapy in Autism Spectrum Disorder: A Systematic Review. Review Journal of Autism and Developmental Disorders, 9, 91–107. https://doi.org/10.1007/s40489-021-00246-x

Ning, Z., Chen, H., Jiang, Y., Hao, C., Ma, G., Wang, S., Yao, J., and Xie, L. (2025). DiffRhythm: Blazingly Fast and Embarrassingly Simple end-to-end Full-Length Song Generation with Latent Diffusion (arXiv:2503.01183). arXiv. https://arxiv.org/abs/2503.01183

Ruth, N., and Müllensiefen, D. (2021). Survival of Musical Activities: When Do Young People Stop Making Music? PLOS ONE, 16, e0259105. https://doi.org/10.1371/journal.pone.0259105

Schneider, F., Kamal, O., Jin, Z., and Schölkopf, B. (2023). Moûsai: Text-To-Music Generation with Long-Context Latent Diffusion (arXiv:2301.11757). arXiv. https://arxiv.org/abs/2301.11757

Wei, J., Karuppiah, M., and Prathik, A. (2022). College Music Education and Teaching Based on AI Techniques. Computers and Electrical Engineering, 100, 107851. https://doi.org/10.1016/j.compeleceng.2022.107851

Wu, Q. (2025). The Application of Artificial Intelligence in Music Education Management: Opportunities and Challenges. Journal of Computational Methods In Sciences And Engineering, 25(3), 2836–2848.

Yu, J., Wu, S., Lu, G., Li, Z., Zhou, L., and Zhang, K. (2024). Suno: Potential, Prospects, and Trends. Frontiers of Information Technology and Electronic Engineering, 25, 1025–1030. https://doi.org/10.1631/FITEE.2400299