Vol.2 No.1 February 2025: 9-14
International Journal of Computer Science and Humanitarian AI (IJCSHAI)
e-ISSN: 3064-4372
DOI: 10.21512/ijcshaijournal.v2i1.13019

Systematic Literature Review of The Use of Music Information Retrieval in Music Genre Classification

M. Aqila Budyputra[1*], Achmad Reyfanza[2], Alexander Agung Santoso Gunawan[3], Muhammad Edo Syahputra[4]

[1-4]Computer Science Department, School of Computer Science,
Bina Nusantara University,
Jakarta, Indonesia 11480
m.budyputra@binus.ac.id; achmad.reyfanza@binus.ac.id; aagung@binus.edu
muhammad.syahputra002@binus.ac.id

***Abstract*— Emphasizing deep learning models such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), this article explores the application of Music Information Retrieval (MIR) techniques in music genre categorization. These algorithms outperform traditional methods in capturing complex audio patterns, showcasing their potential in advancing music classification tasks. Accurate genre classification critically depends on key features such as spectral, temporal, and timbral characteristics, which play a pivotal role in distinguishing musical styles. However, the performance of these models is heavily influenced by the quality and diversity of the training datasets. Additionally, challenges like model interpretability and reliance on large datasets are addressed. This research utilized a Systematic Literature Review (SLR) to investigate the capabilities of advanced MIR techniques in enhancing music categorization systems, particularly for educational applications and personalized music recommendations. The findings reveal that analyzing the importance of spectral, temporal, and timbral features—key components of MIR—can significantly boost the accuracy and reliability of music genre classification.**

Keywords*— *Music Information Retrieval, Music Genre Classification, Deep Learning, CNN, RNN.

I. INTRODUCTION

Music information retrieval (MIR) analysis is a progressively important topic in the modern music scene. As technology advances, MIR is more crucial in determining musical genres and styles, particularly in the fields of digital signal processing and artificial intelligence that allow. The topic has evolved to offer tools and methods that allow a complete grasp of how computer systems classify and analyze music. From music recommendations to plagiarism detection, this evolution has opened doors to various applications that have grown indispensable in the music business and personal listening experiences[1], [2], [3].

Comprising various elements like melody, rhythm, harmony, and texture, music is a worldwide language that accentuates the unique character of every work. Strong awareness of these structures and qualities in the framework of MIR analysis allows us to build computational models capable of spotting and investigating musical trends[4]. With sophisticated algorithms, these models pull out and examine the several components of musical pieces. By doing this, they provide the music industry with significant contributions and generate opportunities for research and closer knowledge of musical compositions. This deeper understanding reaches educational aims, music therapy, and the development of user experience in music streaming services, so boosting the cultural and social impact of music[5], [6]. Particularly MIR research has shown rapid increase. Thanks to significant part to advances in machine learning techniques, especially artificial neural networks—which help systems learn and understand musical features with increased accuracy and depth[7]. By using large and diverse datasets, researchers aim to identify the most accurate and efficient models for musical genre and style detection. Often spanning millions of songs, these databases provide a large reservoir of data enhancing MIR system dependability. Convolutional neural networks (CNNs), recurrent neural networks (RNNs), and other deep learning architectures have greatly improved the precision of genre categorization, mood detection, even musical trend prediction[8].

Combining MIR with other technical innovations such big data analytics and natural language processing (NLP) has also helped to boost its capacities. More sophisticated musical analysis and understanding made possible by these transdisciplinary approaches by mixing words with aural elements, MIR systems can, for example, gather a more whole knowledge of the emotional and contextual background of a song. This all-encompassing approach not only improves the accuracy of music classification but also strengthens customized music recommendation systems, so it is more sensitive to special listener preferences[5], [8].

MIR has great future prospects depending on ongoing research focusing on improving model interpretability, reducing computational costs, and enhancing real-time

Copyright © 2025

processing capabilities[7]. MIR will definitely become ever more crucial in guiding the course of the music business as it grows in providing new means to enjoy, study, and value music. The continuous feedback loop between technical innovation and musical production guarantees that MIR will remain at the forefront of both scientific study and pragmatic application in the music business[3].

Among MIR's options for real-time music analysis and participation, one of the most fascinating ones is MIR systems could offer real-time feedback and modifications depending on the music being performed, therefore altering live events. Furthermore, by raising their responsiveness and engaging power, real-time music analysis could improve interactive music applications including games and educational aids[3]. Using MIR for music therapy offers still another avenue for possible development. By properly identifying and proposing music according on emotional content, MIR could be used to develop tailored treatment sessions that satisfy the emotional needs of individuals[3], [9].

Finally, the way Music Information Retrieval is developing is changing the interaction and viewpoint with music. MIR is making incredible development in many different domains, from enhancing the user experience in music streaming services to raising the accuracy of musical genre classification. Thanks to ongoing research and development, MIR seems brilliant and full of more discoveries that would raise our knowledge and respect of music. MIR's relationship with other emerging technologies will most surely lead to even more creative uses, therefore verifying its indispensible relevance in the music business and outside[5], [10].

II. Methodology

A. Planning the Review

The method used as a preference in conducting research and writing is the study literature review method. This method entails systematically searching, selecting, analyzing, and synthesizing information from diverse literature sources to address research questions or gain a thorough understanding of a topic.

This research finds the systematic literature review method effective for our material evaluation due to its capability to offer a comprehensive grasp of the subject under investigation. By employing this method, the study can gather information and sample data from various pertinent literature sources, analyze and synthesize existing findings, and pinpoint areas where further research is warranted.

B. Research Questions

This research has designed research questions which are the main point in this research as follows:

RQ1: What are the most effective Music Information Retrieval techniques for music genre classification, and how do their accuracies compare?

RQ2: Which features extracted through Music Information Retrieval contribute most significantly to the accuracy of music genre classification?

RQ3: How do the performance and reliability of genre classification models vary when using different datasets and Music Information Retrieval methods?

C. Data Analysis Technique

The data analysis technique this research use in analyzing data with descriptive analysis method is by describing and interpreting the data obtained. This research selects and analyze data or papers obtained using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) method. With the PRISMA method, this research performs identification, systematic review, and selection of relevant literature.

D. Steps of Literature Search

This research is able to conduct a comprehensive and systematic search of a variety of sources of information, including books, research reports, journal articles, and other relevant sources, that are relevant to the subject area, using the literature review method. In order to consolidate and organize pertinent literature and to conduct a critical analysis of each identified source, this study implements a structured methodology. The literature search and review are conducted by adhering to the subsequent procedures:

- Determining the Research Topic

This research initially establishes the subject matter that will be investigated prior to conducting literature research. Following numerous discussions and deliberations, this research has determined to concentrate on the subject of "Systematic Literature Review of The Use of Music Information Retrieval in Music Genre Classification.

- Determining Keywords

In this literature review study, this research uses keywords to facilitate the search for journals that are relevant to the research objectives. This research gathers and selects journals from Google Scholar, PubMed, ScienceDirects, and IEEE Xplore. However, this research mostly obtained them from Google Scholar because it provides mostly free access to journals. Some examples of keywords used are as follows:

TABLE I. Search Terms & Locations

Search locations	Search terms
IEEE Xplore PubMed Google Scholar	(“music” AND “information” AND “retrieval”) OR (“music” AND “genre” AND “classification”)
ScienceDirect	“music information retrieval” OR“music genre classification” OR“audio signal processing”
	“MIR” OR “CNN” OR “RNN”

To ensure the eligibility of the works selected, we apply the selection criteria in Table II using search termsand search locations in Table I.

TABLE II. SELECTION CRITERIA

Inclusion criteria	Exclusion criteria
Internationally recognized paper.	Papers not written in English.
Paper that uses music information retrieval or convolutional neural network for music genre classification.	Papers do not employ music information retrieval or similar methods.
Paper discussing the performance or impact of music information retrieval.	
Paper discussing about music genre classification.	
Paper is published within the last 5 years (With exception for papers with very high relevance and contributes a lot to the topic).	

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) Flow Chart The collection, selection, and analysis of data in this study use the PRISMA

(Preferred Reporting Items for Systematic Reviews and Meta-Analysis) method. The selection process is divided into, Screening, Eligibility, and Included. The data collection using the PRISMA method is divided into four stages, which are:

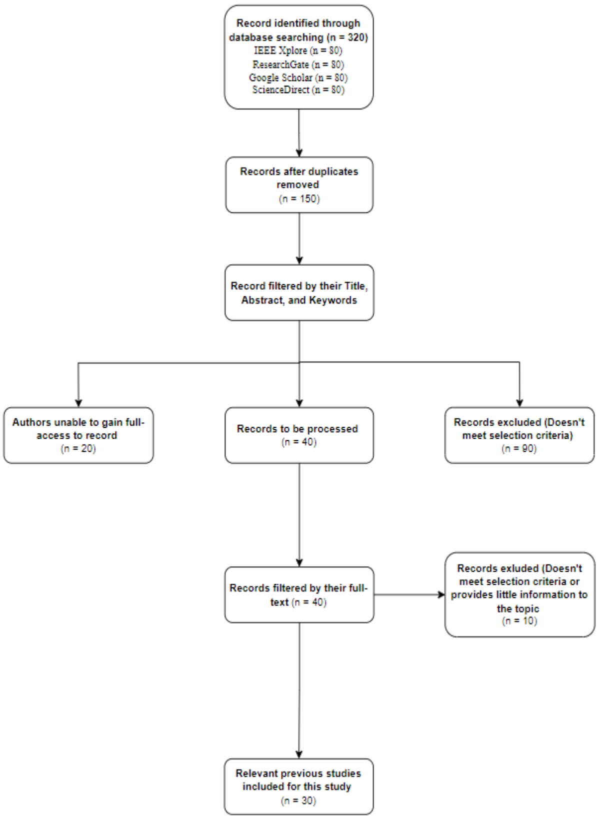


Fig. 1. PRISMA Flowchart

III. RESULT

The results of this paper are the answers to the research questions mentioned earlier. The answers are based on a literature study that involved searching through various academic sources such as IEEE Xplore, Google Scholar, and others. From this search, a total of 30 relevant papers.

RQ1: What are the most effective Music Information Retrieval techniques for music genre classification, and how do their accuracies compare?

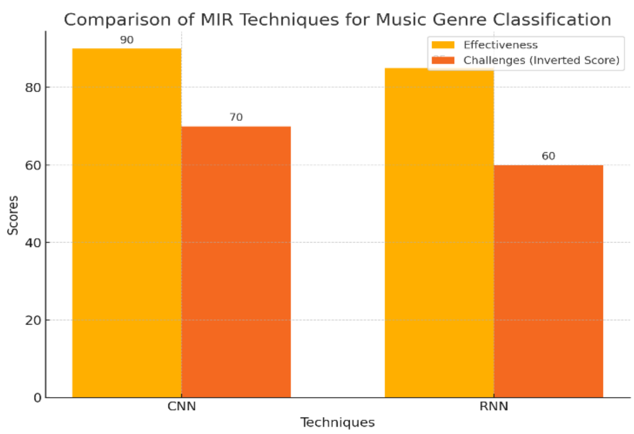
Based on Music Information Retrieval (MIR) approaches, deep learning models—such as CNNs and RNNs—have shown to be quite successful in music genre recognition. CNNs are very successful in studies with amazing results signaling improved accuracy and the capacity to identify complicated patterns in audio data[11], [12], [13]. Though much less than CNNs, RNNs have also shown really high performance[14], [15], [16].

These deep learning models can automatically classify music into different genres and surpass conventional techniques with great accuracy[11], [12], [13]. their capacity to identify complex elements in audio data greatly affects their outstanding performance in genre categorization.[14], [15], [17].

Still, the problems deep learning approaches bring in this field are really fascinating. These elements clearly relate to the quality and quantity of training data, the complexity of applying these models, and the challenge of analyzing their outputs[11], [12], [14], [18]. Among the many difficulties maximizing the possibilities of deep learning models in music genre categorization needs addressing assuring the availability of excellent data, fine-tuning model architecture, and enhancing result interpretation[13], [19], [20].

Using more varied datasets lets data augmentation methods include altering pitch and speed and combining audio sections improve the classification performance. Still, several research have shown that the success of these strategies can have certain restrictions[15].

CHART I. COMPARISON OF CNN AND RNN TECHNIQUES



RQ2: Which features extracted through Music Information Retrieval contribute most significantly to the accuracy of music genre classification?

The MIR aspects that best capture the spectral, temporal, and timbral characteristics of the audio are those that help to differentiate between many music genres. The classification method examines the harmonic and pitch content of the music using spectral factors like chroma. This helps the program to spot unique trends in the harmonic progressions unique to several genres. Crucially in deciding the genre, these qualities provide a clear picture of the harmonic structure of the song[14], [15].

Mostly depending on temporal data, including beat-related statistics and rhythmic patterns, music genre categorization is These components cover the song's rhythmic qualities and may show clear differences across many genres. They also are rather important in producing the particular atmosphere of every musical genre[12], [13].

Moreover, it has been shown that the timbral qualities of the music help to differentiate between other genres. Different instruments, manufacturing methods, and sound qualities connected with different musical genres depend critically on timbral aspects like spectral centroid, rolloff, and flux[11], [12].

The relevance of various kinds of characteristics might change depending on the genres under analysis, the dataset used, and the particular categorization techniques used. Carefully engineering and selecting suitable characteristics will help to improve the effectiveness of genre classification systems for a given use[11], [15].

Including spectral, temporal, and timbral data helps music genre classification algorithms better discriminate between different musical genres and categories. Further research in this field might help to improve the reliability and precision of automated music genre recognition systems[12], [14].

TABLE III. FEATURES IMPACT FOR GENRE CLASSIFICATION

Feature Type	Description	Genre Classification Impact	Key Refere-nces
Spectral	Chroma, harmonic, and pitch content, identifies patterns in harmonic progression	Provides compact representation of music's harmonic structure	[14], [15]
Temporal	Beat-related statisticsm rhythmic properties of music	Varies significantly across genres, contributes to the unique feel of a music style	[12], [13]
Timbral	Spectra centroid, rolloff, flux, describes sonic quality or color of the sound	Helps identify specific instruments, production techniques, and sonic signatures	[11], [12]

RQ3: How do the performance and reliability of genre classification models vary when using different datasets and Music Information Retrieval methods?

The performance and dependability of music genre classification models may be much changed by the employed datasets and Music Information Retrieval (MIR) techniques[12], [13], [21], [22]. Usually, models combining spectral, temporal, and timbral characteristics outperform those depending just on one kind of feature[11], [12], [23].

Both deep learning general and contemporary MIR techniques have greatly improved genre categorization accuracy. Particularly deep learning methods enable detect complex patterns in audio data, thus generating suitable genre classifications[11], [12], [13], [14], [22], [24]. Still, factors like dataset bias and the difficulty evaluating the results of complex systems might compromise the reliability of these models.[11], [14], [25], [26].

Examining genre categorization models closely across several datasets and genres helps to guarantee generalizability. Analyzing genre categorization models across many datasets and genres enables one to ensure generalizability. The optimum approach is to choose representative, high-quality datasets[11], [12], [14] and use well-crafted MIR features and modeling algorithms appropriate for the individual application demands. Dealing with issues like the complexity of model implementation, the necessity of high-quality training data, and the difficulties linked with result interpretation will assist to enhance the performance of deep learning models in music genre classification[11], [13].

Fundamentally, the efficiency and dependability of genre classification models control the choice of datasets, the MIR algorithms used, and the ability to solve the associated challenges. Constant research and development, especially in the application of deep learning to tackle its shortcomings[12], [14], [27] will yield more powerful and consistent responses for music genre classification.

TABLE IV. IMPACT OF VARIOUS FACTORS ON THE PERFORMANCE AND RELIABILITY OF GENRE CLASSIFICATION

Aspect	Impact on Model Performance	Key References
Dataset Influence	Comprehensive genre coverage and clear distinctions between genres lead to higher accuracy	[11], [12], [13], [21], [22], [24]
Feature Utilization	Combination of spectral, temporal, and timbral feautres outperforms single-feature approaches	[11], [12], [22], [23], [28], [29]
MIR Techniques	Deep learning models recognize complex patterns effectively, enhancing classification accuracy	[11], [12], [13], [14], [22], [24]
Model Reliability	Reliability may be compromised by dataset bias and complexity in model interpretation	[11], [14], [25], [26], [30]
Generalizability	Requires careful evaluation across diverse datasets to assess and ensure model generalizability	[11], [12], [14], [27]

IV. CONCLUSION

Deep learning methods have fundamentally enhanced the performance of automated music genre classification systems. These sophisticated models find complex trends in audio data, therefore surpassing conventional techniques. Still, the quality and kind of training data will mostly determine how successful these models are.

The reliance on accessible datasets in this discipline presents one of its key difficulties. The degree of diversity, representativeness, and excellent quality of data determines directly how well these model's function. Better accuracy usually comes from datasets spanning a wide spectrum of genres and with obvious differences between them. This is so because deep learning models require vast and varied datasets to properly grasp the intricate characteristics distinguishing different music genres.

Furthermore, challenging to understand which particular features contribute most to the accuracy of deep learning models is their complexity. This calls for concentrated work on improving models as well as data. Analyzing spectral, temporal, and timbral properties among common Music Information Retrieval (MIR) feature extraction techniques shows different degrees of significance in improving the accuracy of genre classification.

Future studies should aim at creating standardized, high-quality datasets that better reflect the variety of music genres in order to meet these problems. Furthermore, efforts should be made to make deep learning models more interpretable so that one may better grasp the essential characteristics and trends guiding correct genre classification. Researchers and practitioners can make major progress in the automated music genre classification by using deep learning together with thorough data analysis and MIR techniques. More solid and dependable systems for use including music recommendation, music information retrieval, and music education could follow from this advancement. These developments greatly suggest the possibility for more precise and perceptive systems of music genre classification.

References

[1] by Y. V Srinivasa Murthy, "Content-based Music Information Retrieval (CB-MIR) and its Applications Towards Music Recommender System," 2018.

[2] R. Gupta, J. Yadav, and C. Kapoor, "Music Information Retrieval and Intelligent Genre Classification," 2021, pp. 207–224. doi: 10.1007/978-981-15-8443-5_17.

[3] P. Kasák, R. Jarina, and M. Chmulik, "Music information retrieval for educational purposes - an overview," Mar. 2020, pp. 296–304. doi: 10.1109/ICETA51985.2020.9379216.

[4] S. O. Folorunso and F. E. Ayo, "Machine Learning Analysis of Music Based on Music Information Retrieval Tasks," 2022. doi: 10.21467/proceedings.157.

[5] M. Schedl, E. Gómez, and J. Urbano, "Music Information Retrieval: Recent Developments and Applications," *Foundations and Trends® in Information Retrieval*, vol. 8, no. 2–3, pp. 127–261, 2014, doi: 10.1561/1500000042.

[6] M. Gaston Michel Albert Devaux and N. Thanh Tam, "Music Information Retrieval with Deep Sequential Models," 2019, doi: 10.13140/RG.2.2.33458.45761.

[7] H. Purwins, B. Li, T. Virtanen, J. Schlüter, S.-Y. Chang, and T. Sainath, "Deep Learning for Audio Signal Processing," *IEEE Journal on Selected Topics in Signal Processing*, vol. 13, no. 2, pp. 206–219, 2019, doi: 10.1109/JSTSP.2019.2908700.

[8] J. Ramírez and M. J. Flores, "Machine learning for music genre: multifaceted review and experimentation with audioset," *J Intell Inf Syst*, vol. 55, no. 3, pp. 469–499, 2020, doi: 10.1007/s10844-019-00582-9.

[9] B. Kostek, "Music information retrieval - The impact of technology in art, crowdsourcing, big data, and the cloud," in *Proceedings of Meetings on Acoustics*, Acoustical Society of America, 2019. doi: 10.1121/2.0001225.

[10] Y. Hu and G. Mogos, "Music genres classification by deep learning," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 25, no. 2, pp. 1186–1198, Feb. 2022, doi: 10.11591/ijeecs.v25.i2.pp1186-1198.

[11] Z. Nasrullah and Y. Zhao, "Music Artist Classification with Convolutional Recurrent Neural Networks," in *Proceedings of the International Joint Conference on Neural Networks*, Institute of Electrical and Electronics Engineers Inc., Jul. 2019. doi: 10.1109/IJCNN.2019.8851988.

[12] T. Ziemer, P. Kiattipadungkul, and T. Karuchit, "Acoustic features from the recording studio for music information retrieval tasks," in *Proceedings of Meetings on Acoustics*, Acoustical Society of America, 2020. doi: 10.1121/2.0001363.

[13] N. Choudhury, D. Deka, S. Sarmah, and P. Sarma, "Music Genre Classification Using Convolutional Neural Network," in *2023 4th International Conference on Computing and Communication Systems, I3CS 2023*, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/I3CS58314.2023.10127554.

[14] M. Schedl, "Deep Learning in Music Recommendation Systems," *Frontiers in Applied Mathematics and Statistics*, vol. 5. Frontiers Media S.A., Aug. 29, 2019. doi: 10.3389/fams.2019.00044.

[15] A. Kratimenos, K. Avramidis, C. Garoufis, A. Zlatintsi, and P. Maragos, "Augmentation methods on monophonic audio for instrument classification in polyphonic music," in *European Signal Processing Conference*, European Signal Processing Conference, EUSIPCO, Jan. 2021, pp. 156–160. doi: 10.23919/Eusipco47968.2020.9287745.

[16] D. Mi and L. Qin, "Classification System of National Music Rhythm Spectrogram Based on Biological Neural Network," *Comput Intell Neurosci*, vol. 2022, 2022, doi: 10.1155/2022/2047576.

[17] T. Li, "Optimizing the configuration of deep learning models for music genre classification," *Heliyon*, vol. 10, no. 2, p. e24892, 2024, doi: https://doi.org/10.1016/j.heliyon.2024.e24892.

[18] P. Patel, "Music Retrieval System Using Query-by-Humming," San Jose State University, San Jose, CA, USA, 2019. doi: 10.31979/etd.mh97-77wx.

[19] B. Wilkes, I. Vatolkin, and H. Müller, "Statistical and Visual Analysis of Audio, Text, and Image Features for Multi-Modal Music Genre Recognition," *Entropy*, vol. 23, no. 11, 2021, doi: 10.3390/e23111502.

[20] H. Li, X. Fei, and K.-M. Chao, "E3MSD-A New Music Information Retrieval Architecture for an Original Music Identifier," 2020. [Online]. Available: https://www.researchgate.net/publication/344658704

[21] Y. Chu, "Recognition of musical beat and style and applications in interactive humanoid robot," *Front Neurorobot*, vol. 16, Jul. 2022, doi: 10.3389/fnbot.2022.875058.

[22] X. Jia, "A Music Emotion Classification Model Based on the Improved Convolutional Neural Network," *Comput Intell Neurosci*, vol. 2022, pp. 1–11, Jul. 2022, doi: 10.1155/2022/6749622.

[23] N. Kumari, "Music Genre Classification for Indian Music Genres," *Int J Res Appl Sci Eng Technol*, vol. 9, no. 8, pp. 1756–1762, Aug. 2021, doi: 10.22214/ijraset.2021.37669.

[24] E. Y. Koh, K. W. Cheuk, K. Y. Heung, K. R. Agres, and D. Herremans, "MERP: A Music Dataset with Emotion Ratings and Raters' Profile Information," *Sensors*, vol. 23, no. 1, 2023, doi: 10.3390/s23010382.

[25] X. Chen, X. Qu, Y. Qian, and Y. Zhang, "Music Recognition Using Blockchain Technology and Deep Learning," *Comput Intell Neurosci*, vol. 2022, no. 1, p. 7025338, 2022, doi: https://doi.org/10.1155/2022/7025338.

[26] Q. Lin, "Music Score Recognition Method Based on Deep Learning," *Comput Intell Neurosci*, vol. 2022, Jul. 2022, doi: 10.1155/2022/3022767.

[27] X. Jia, "A Music Emotion Classification Model Based on the Improved Convolutional Neural Network," *Comput Intell Neurosci*, vol. 2022, 2022, doi: 10.1155/2022/6749622.

[28] K. Pyrovolakis, P. Tzouveli, and G. Stamou, "Multi-Modal Song Mood Detection with Deep Learning," *Sensors*, vol. 22, no. 3, 2022, doi: 10.3390/s22031065.

[29] F. Simonetta, S. Ntalampiras, and F. Avanzini, “Multimodal music information processing and retrieval: Survey and future challenges,” in *Proceedings - 2019 International Workshop on Multilayer Music Representation and Processing, MMRP 2019*, Institute of Electrical and Electronics Engineers Inc., Mar. 2019, pp. 10–18. doi: 10.1109/MMRP.2019.8665366.

[30] M. Ohrnberger, “From Music Information Retrieval to Innovative Seismological Signal Processing Schemes,” 2021. [Online]. Available: https://www.researchgate.net/publication/357339043