

Project Name - Adapting music to dramatic scenes in movies

Group Partners - Kim Tsadok 206821084

Background –

We all know and love movies, and TV shows, they surround all of us on a daily basis. from all different Genres like horror, adventure, thriller, romance, comedy and so on.

One of the main things that get most people hooked on a certain Tv show or movie is the music behind the scenes.

Music that cater to each vibe perfectly, further improving and enhancing the creator's meaning behind the scenes.

We plan to help all those looking to create content on TV and cinemas with their search for the perfect song or music.

We wish to develop this system in order to ease on these film makers, theatre students, directors, producers, videographers, etc. with their quest to further emphasize their creation on the big screen.

Our Goals –

Our goal with this project is to develop an AI based system with Machine Learning that will analyze Dramatic scenes and scenarios from movies, shows, and so on, that will cater a specific audio to enhance and further improve the emotional significance of the scene, to its needs.

Our Objectives –

- Developing a user-friendly system with a comfortable UI to help content creators, including film students, producers, and independent video makers, access and navigate in it easily to their needs
- Develop an insert tool to a specific clip for a movie, show, etc. for our model to cater a specific music to.
- Develop an option for tags for the model to more specifically feature a more proper music for the specific scene.
- Compile the collected data into a database for the model to rely on and further improve as we go on the project.

Literature Review

Artificial Intelligence (AI) has revolutionized various creative domains, including music recommendation and film production. The integration of machine learning techniques in analysing multimedia content has gained significant traction in recent years. This literature review explores key studies and methodologies relevant to developing an AI-based system that analyses dramatic scenes in films and suggests emotionally appropriate music.

AI in Music Recommendation

Music recommendation systems traditionally rely on collaborative filtering, content-based filtering, or hybrid approaches. Collaborative filtering predicts user preferences based on similarities between users, while content-based filtering focuses on the attributes of the music itself, such as tempo, genre, and mood. Recent advances have seen the adoption of deep learning models for more nuanced music analysis, leveraging embeddings to capture complex patterns in audio features (Chowdhury et al., 2020). However, these systems often lack contextual awareness of multimedia environments, making them less effective for scene-specific recommendations. (such as drama scenes in movies, etc)

Visual and Audio Analysis in Multimedia

The use of AI for scene analysis in films typically involves computer vision and audio processing techniques. Computer vision models, particularly convolutional neural networks (CNNs), have demonstrated exceptional capabilities in recognizing visual features such as lighting, color palettes, and camera movement (Kaur et al., 2019). Similarly, recurrent neural networks (RNNs) and transformers are employed to process audio data, capturing nuances in dialogues, sound effects, and background scores. Integrating visual and audio modalities for scene analysis presents a promising approach to understanding the emotional context of cinematic content (Zhao C Liu, 2021).

Emotional Context and AI

Emotional analysis has been a growing area of research, with applications in sentiment analysis and affective computing. Studies like Wang et al. (2022) emphasize the importance of multimodal emotion recognition, combining facial expressions, audio tone, and textual cues. This approach aligns with the objectives of our project, which seeks to enhance the emotional resonance of scenes through music. Existing research underscores the need for precise emotion recognition to achieve seamless multimedia integration. (Like what we plan to do in our project through Machine learning and AI).

Gaps in Current Systems

While there are systems that offer general music recommendations based on user input or predefined parameters, few incorporate deep multimodal analysis for scene-specific emotional matching. This gap highlights the potential for innovation in integrating AI-driven image and sound analysis with music curation, as explored in this project.

Competition Review

The market for AI-driven multimedia tools is expanding rapidly, with several solutions addressing aspects of music recommendation and video editing. This section reviews existing tools and systems, comparing their capabilities with the proposed project.

Existing Music Recommendation Systems

Spotify and Apple Music dominate the commercial music recommendation space, leveraging sophisticated algorithms to personalize user playlists. While effective for general music discovery, these platforms do not cater to the nuanced requirements of filmmakers, such as scene-specific music suggestions. Similarly, tools like Amper Music and AIVA provide AI-generated compositions, focusing on music creation rather than selection.

Scene Analysis Tools

Adobe Premiere Pro and Final Cut Pro offer advanced editing features, including basic scene detection. However, their capabilities are limited to identifying cuts and transitions without delving into emotional context. Tools like Runway ML have introduced AI-driven enhancements in video editing, such as object recognition and background replacement, but lack integration with music selection processes.

Emerging Innovations

Startups like Endel and Melodrive are exploring AI in adaptive music generation. Endel uses environmental cues to create personalized soundscapes, while Melodrive focuses on interactive music generation for virtual environments. Although innovative, these solutions primarily cater to immersive experiences rather than cinematic storytelling.

Competitive Advantage

The proposed system (Our project) stands out by combining deep multimodal analysis—visual and audio—to understand the emotional essence of a scene. Unlike existing tools, our project focuses on aligning pre-existing music tracks to cinematic content, streamlining the workflow for filmmakers. This approach reduces the dependency on musical expertise, providing a unique value proposition to content creators, including film students, producers, and independent video makers.

By addressing the limitations of current tools and building upon advances in AI and multimedia analysis, our project aims to fill a critical gap in the creative industry.

References

1. Runway ML's AI-based video editing tools, which focus on scene analysis [Runway ML](#).
2. *Artificial Intelligence and Music Ecosystem* by Martin Clancy, 2022. Focuses on AI's role in music creation and recommendation.
3. Zhao, Y., C Liu, K. (2021). *Multimodal Scene Analysis Using Deep Learning*. Journal of Multimedia Processing.
4. Wang, T., et al. (2022). *Emotion Recognition in Multimedia Contexts*. Proceedings of the International Conference on Affective Computing.
5. Chowdhury, S., et al. (2020). *Improving Music Recommendations with Deep Learning*. Journal of Artificial Intelligence Research.
6. Kaur, N., et al. (2019). *Computer Vision Techniques in Film Scene Analysis*. International Journal of Computer Vision.
7. IEEE. (2020). "Deep Learning in Intelligent Music Recommendation Systems." *IEEE Xplore*. Available at: <https://ieeexplore.ieee.org/document/9319393>