DEEP LEARNING-BASED AUTOMATIC MUSIC TRANSCRIPTION USING CR-GCN

126003238 Selvakarthik S
126003241 Shanthosh Kumar R
126003218 Rithvik L

ABSTRACT:

The task of automatic music transcription (AMT) mainly focuses on converting audio signals to symbolic music representations, facilitating applications such as computational musicology and music analysis. One of the biggest problems is when multiple notes are played at the same time, dimension explosion can happen which makes it difficult for accurate music note transcription. To overcome this challenge, we have proposed a hybrid deep learning architecture combining Convolutional Neural Network for spatial feature extraction, bidirectional LSTMs or self-attention mechanisms for precise temporal note-level predictions and Graph Convolutional Network for accurate label learning to capture note interdependencies in polyphonic music. Experiments on public datasets like MAESTRO, MAPS, GiantMIDI show that the proposed methodology with F1-score of 96.88% is much more superior than existing methodologies like Onset and Frames, Wavenet, Non-Negative Matrix Factorization (NMF). The generated music sheets validate the model's accuracy and practical applicability, providing a valuable tool for musicians and researchers. By addressing the limitations of prior methods, the proposed approach CR-GCN (Channel Relationship-Based Graph Convolutional Network) represents a step forward in automated transcription technology, making it feasible for large-scale and real-time applications.

References:

[1] Xiao, Z., Chen, X., & Zhou, L. (2023). Polyphonic piano transcription based on graph convolutional network. *Signal Processing*, *212*, 109134.

Dr. A. Emily Jenifer