**Introduction**

Music classification has always been at the core of music information retrieval systems. Till now, music genres have been classified and labelled into classes manually. This is a very time consuming and tedious process. This was also necessary because of the evolving nature of music. The boundary between different genres has been blurred as Artists of one genre have been influenced by artists of other genres. However, it has been observed that audio signals (digital or analogue) of music belonging to the same genre share certain characteristics, because they are composed of similar types of instruments, having similar rhythmic patterns, and similar pitch distributions. This suggests feasibility of automatic musical genre classification.

**Literature Survey**

CNN’s are used widely for image processing and processing extracted features from audio clips. Pons et, al. [1] propose using wider and a greater number of filters and 2 parallel O-net and P-net CNNs in the first layer of the CNN for better feature extraction more suitable for music to extract features related to rhythm, tempo etc. Features extracted from the CNN can also be passed in to an RNN [2] which, due to its nature preserves temporal and local characteristics and is able to give more “attention” to the local features helpful in classification. Another approach used by Tao Li et, al. [3] is by using the Daubechies Wavelet Coefficient Histograms (DWCH) wavelet transforms for feature extraction and classification. Mohsin Ashraf et, al. [4] discuss the “batch normalisation” problem due to fixed batch sizes in the CRNN architecture and propose a Globally Regularised Neural Network which produce satisfactory results. Tao Feng [5] has proposed a method where they pre-train RBMs (Restricted Boltzmann machines) iteratively between layers and stack them together on the multilayer architecture in the end. This essentially becomes a Deep Belief Network (DBN), which gave better results on larger datasets. D Ghosal et al. [6] have proposed a method where Conventional recurrent neural network approach was used. This approach uses RNN on top of CNN which is used as feature extractor. The output of this is passed onto a classifier, which classifies based on genre. They have sent multiple frames of the song as input which gave better results compared to a single frame. Mingwen Dong [7] has given an approach in which we use CNN to extract features from mel-spectrograms of audio. On further inspection, the features learned by the CNN in 2nd max pooling layer was found to be similar to the spectro-temporal receptive field observed in the human brain, concluding this architecture learns similar to the human brain. Hansi Yang, Wei-Qiang Zhang [8] have proposed a method which uses duplicate convolutional networks. They duplicate 3 and 4 CNNs to check the results. The output from one CNN is concatenated to the output of the next CNN. This output is then fed into a classifier for classification.

Weibin Zhang et, al[9] proposed two ways to improve music genre classification with convolutional neural networks (CNN): 1) combining max- and average pooling to provide more statistical information to higher level neural networks; 2) using shortcut connections to skip one or more layers, a method inspired by residual learning method. GTZAN dataset consisting of about a thousand songs and 10 major genres is used, accuracy of 84.8% and 87.4% is observed respectively. Sergio Oramas et, al[10] propose to categorize musical items into multiple and fine-grained labels, using three different data modalities: audio, text and images using optimal models for each modals. Alexandros Tsaptisinos[11] propose lyrics based genre classification using Hierarchical Attention Network (HAN) also HAN is compared with many baseline models. Hareesh Bahuleyan[12] proposed an ensemble classifier combining VGG-16 CNN and Extreme Gradient Boosting, from spectrogram of the song as input, the results showed AUC value of .894. Trained using *Audio Set* dataset.

**References**

[1] - J. Pons and X. Serra, "Designing efficient architectures for modeling temporal features with convolutional neural networks," 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2017, pp. 2472-2476, doi: 10.1109/ICASSP.2017.7952601.

[2] - K. Choi, G. Fazekas, M. Sandler and K. Cho, "Convolutional recurrent neural networks for music classification," 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2017, pp. 2392-2396, doi: 10.1109/ICASSP.2017.7952585.

[3] - Li, Tao & Ogihara, Mitsunori & Li, Qi. (2003). A Comparative Study on Content-Based Music Genre Classification. 282-289. 10.1145/860435.860487.

[4] - M. Ashraf, G. Geng, X. Wang, F. Ahmad and F. Abid, "A Globally Regularized Joint Neural Architecture for Music Classification," in IEEE Access, vol. 8, pp. 220980-220989, 2020, doi: 10.1109/ACCESS.2020.3043142.

[5] - Feng, T., 2014. Deep learning for music genre classification. *private document*.

[6] - Ghosal, Deepanway, and Maheshkumar H. Kolekar. "Music Genre Recognition Using Deep Neural Networks and Transfer Learning." In *Interspeech*, pp. 2087-2091. 2018.

[7] - Dong, Mingwen. "Convolutional neural network achieves human-level accuracy in music genre classification." *arXiv preprint arXiv:1802.09697* (2018).

[8] - Yang, Hansi, and Wei-Qiang Zhang. "Music Genre Classification Using Duplicated Convolutional Layers in Neural Networks." In *INTERSPEECH*, pp. 3382-3386. 2019.

[9] W. Zhang, W. Lei, X. Xu, and X. Xing, “Improved music genre classification with convolutional neural networks.,” in INTERSPEECH, pp. 3304–3308, 2016.

[10] Sergio Oramas, Oriol Nieto, Francesco Barbieri, and Xavier Serra. Multi-label music genre classification from audio, text, and images using deep features. International Society for Music Information Retrieval Conference (ISMIR), 2017.

[11] A. Tsaptsinos, “Lyrics-based music genre classification using a hierarchical attention network,” CoRR, vol. abs/1707.04678, 2017.

[12] H. Bahuleyan, “Music genre classification using machine learning techniques,” arXiv preprint arXiv:1804.01149, 2018.