IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.918 ⁜ Vol. 12, Issue 2, February 2023

DOI: 10.17148/IJARCCE.2023.12231

Music Genre Classification

Prof. Manjunath T N[1], R Kavana[2], Raksha M K[3]

Associate Professor, Dept of Computer Science, East West Institute of Technology, Bangalore, India[1]

UG Student, Dept of Computer Science, East West Institute of Technology, Bangalore, India[2]

UG Student, Dept of Computer Science, East West Institute of Technology, Bangalore, India[3]

Abstract: Music classification is a field in the field of Music Recovery and sound signal processing research. Neural Network is a modern way of classifying music. The classification of music using neural networks (NN) has become very successful in past few years. Different song collections, machine learning methods, input formats, and neural network applications are all to varying degrees effective. Spectrograms made from time-slices of songs are input to a neural network in order to classify songs into the appropriate musical genres. The Neural Network (NN) employs spectrograms generated by time song slaves as an entry to classify songs into their numerous genres. The Convolutional Neural Network (CNN) audio signal input system will employ the generated spectrograms. Tasks involving picture pattern recognition are handled by CNNs. Acoustic feature extraction is the most important process while evaluating music. Models are trained using the GTZAN dataset in the suggested system.

Keywords: Deep learning, spectrogram, music, classification of music genre, Convolution Neural Network (CNN).

I. INTRODUCTION

A. Machine Learning

Machine learning is currently trending in the technological arena. According to the requirement of the application and the available dataset different machine learning techniques can be applied to a certain problem but finding the most appropriate algorithm for the scenario is significant. Algorithms for machine learning often fall into one of four groups. The four distinct categories of learning algorithms are reinforcement learning, supervised learning, unsupervised learning, and semi-supervised learning. Applying supervised learning to a dataset that has been entirely tagged will result in the creation of a mathematical model. Regarding unsupervised learning keeping a target in mind it extracts the features of the data in the dataset. Semi-supervised learning is the hybrid learning method which utilizes both labelled and unlabeled datasets. Reinforcement learning determines how an agent automatically acts in the environment to obtain desired outcomes.

B. Neural Networks

A machine learning technique called a neural network (NN) effectively filters important feature elements in big datasets and creates a product or model that captures those attributes. The model is first trained by NN using the training database. After model training, NN can be used to separate data using previously trained models on fresh or previously unselected data points.

CNNs, a subset of machine learning, include convolutional neural networks (CNNs). CNN is a specific kind of deep learning network design that is used for operations like image recognition and pixel-level data processing The four different classifications of learning algorithms are. The foundation of deep learning systems are synthetic neural networks (ANNs). Recurrent neural networks are one ANN method that takes input from time series or sequential data (RNN). Natural language processing (NLP), language translation, speech recognition, and image captioning are a few of its applications.

A network that can find important information in both time series and picture data is CNN, for instance. As a result, it is very helpful for applications like image identification, object classification, and pattern recognition. CNNs use matrix multiplication and other linear algebraic techniques to find patterns in images. Signal and audio data can be categorized by CNNs.

CNN's organizational structure is similar to the way connections are made in the human brain. Similar to the billions of neurons in the human brain, CNNs too feature neurons that are arranged in specific groups. CNN can evolve on current networks and is flexible enough to work in new roles. Using parameter sharing, CNNs are more computationally efficient

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 7.918 ⁜ Vol. 12, Issue 2, February 2023
DOI: 10.17148/IJARCCE.2023.12231

than conventional NNs. Models created by the CNN algorithm are simple to use and functional on all platforms. Convolutional, pooling, and fully linked layers make up the three levels of CNN. Each node on a CNN layer connects to another node that is already there. Healthcare, automotive, social media, retail, and audio processing are a few industries that use CNN.

C. Music Classification

Music classification is a type of music retrieval function (MIR) where labels are assigned to music features such as genre, heart rate, and instruments. It is also associated with concepts such as musical similarities and musical tastes. People have known about music since the beginning of time. The concept of classification of music allows us to distinguish between different genres based on their composition and frequency of hearing.

With the increasing variety of genres around the world, the classification of genres has recently become quite popular. The classification how to improve an intelligence interview agent that can of genres is an important step in building a useful commendation system for this project. The ultimate goal is to create a machine learning model that can more accurately classify music samples into different genres. These audio files will be categorized based on the minimum frequency and time zone characteristics.

II. LITERATURE SURVEY

This section shows the survey done on different papers about the same work area. We'll compare the work that has already been implemented, so that we could improvise the proposed system for better efficiency. The following are some existing system analysis.

Title: Neural Network Genre Classification

With certain modifications, this research employs a neural network that is comparable to Despois's work. Two examples of the changes include an increase in the categories for music genres and the utilization of multiple music databases. The activation function, the number of training spectrogram slices used to train the neural network, and the number of training spectrogram slices per genre all underwent changes. Early statistics decreased the number of genres to seven, which now include Hip-Hop, Electronics, Rock, Pop, R&B, Alternative, and Country. The main 7 genres were combined with a number of sub-genres to achieve this. The collection was roughly doubled and increased by 132,000 spectrogram slices, or 1880 songs. After the changes, a 62% accuracy rate for the test was found. The accuracy was then improved in the convolutional layers by 5% by switching the activation function from an Exponential Linear Unit (ELU) to a ReLU.

Advantages**:** Convolution neural network with a softmax function, and a one-hot array of genre classifications to determine the likelihood that each genre would be recognized..

Disadvantages**:** The initial results on the test data were 67% accurate.

Title: A Study on Broadcast Networks for Music Genre Classification.

The general public has become more interested in Music Genre Classification as a result of the rising use of music streaming and recommendation services, as well as recent advancements in frameworks for music information retrieval (MGC). Contrarily, it is well known that convolutional-based approaches fail miserably at encoding and localizing temporal data. With the goal of enhancing localization and generalizability within a set of restricted parameters, we analyze broadcast-based neural networks in this research and investigate twelve broadcast network versions (about 180k).

We discuss the effects of block configuration, the pooling method, the activation function, the normalizing process, the label smoothing, the interdependency of the channels, the addition of the LSTM block, and other changes of the inception scheme. Using key datasets including GTZAN, Extended Ballroom, HOMBURG, and Free Music Archive, our computational studies demonstrate the cutting-edge classification accuracy of MGC (FMA). With the aid of our method, tiny broadcast networks for music classification could be created that are widely applicable.

Advantages: provides information and the opportunity for generalized, compact broadcast networks for music classification.

Disadvantages: The model's precision could be increased.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 7.918 ⁜ Vol. 12, Issue 2, February 2023
DOI: 10.17148/IJARCCE.2023.12231

Title: Using Deep Learning to Identify Multilingual Music Genre

Nowadays, the topic of automatic genre classification is heavily researched. Given the enormous increase in the amount of music being released, it is a crucial responsibility. According to a poll, Spotify alone puts out around 10,000 songs each month. Users simultaneously acquired a liking for listening to music in several languages. Maintaining a database that can automatically categorize music of any language based on their genres is therefore essential. The goal of this research is to use a deep learning convolutional neural network technique to categorize a song's genre across twelve distinct languages. This study mainly divided the sixty songs in a database into the categories of metal, hip hop, and pop. Each song is cut to exactly five seconds, or a "slice," which is then transformed into a spectrogram—a graphic representation of a song's time-domain spectrum. There are training and testing datasets for these spectrograms. The data set is divided into 25% for testing and 75% for training. To automatically categorize the songs into their many genres, an image classifier called a tensor flow deep learning convolutional neural network is employed.

Advantages: For all twelve languages taken into account in this database, this model's classification accuracy was 93.3%

Disadvantage: Since localization and dilatation cannot be balanced, the model can record both high- and low-frequency components.

Title: Music Genre Recognition Using Residual Neural Network

Despite being abstract, genre is a distinguishing aspect of music. Existing methods for categorizing genres automatically compute several audio features before layering a classifier on top of them. These models frequently perform these calculations over a substantial chunk of the audio. This study suggests a genre categorization model based on residual neural networks that is trained on short (3-second) video snippets. Additionally, conventional genre categorization algorithms will only give an audio clip a single genre. But it's commonly known that numerous genres have certain traits. The methods suggested in this article can assign three genre labels—each related with a specific probability—to a music clip in light of the genre's uncertainty. The suggested model mispredicted the top 1, top 2, and top 3 genres for a music clip with error rates of 18%, 9%, and 5.5%, respectively. Here, we show how the classifier's predictions in a real-world situation correspond to the broader notion of genre.

Advantages**:** An illustrated residual deep learning model for genre classification had accuracy rates of 82%, 91%, and 94.5% for the top-1, top-2, and top-3 most likely genre classes assigned to an audio sample at test time.

Disadvantages: The model's fidelity could be increased.

III. CONCLUSION

The suggested method has examined some machine learning research that has been done on music and genre classification. Additionally, basic study and implementation work for musical genre classification have been presented. Songs were broken down into brief time segments for the research implementation, and the time segments were then represented by the corresponding spectrogram images. A music genre was assigned to each of these spectrograms, and the neural network used those labels as inputs. To calculate the likelihood that each genre would be recognized, the neural network used four convolutional layers, a fully connected layer, a softmax function, and a one-hot array of genre classifications.

IV. ACKNOWLEDGEMENT

We are appreciative to have reached this point in our mission because it required a lot of help and work from many people to finish and had such a big impact on our enterprise. We are immensely appreciative and humbled by everyone who has helped us transform these concepts into something more profound.

We would especially like to express our gratitude to **prof. Manjunath T N** (Professor, Computer Science and Engineering, East West Institute of Technology), who gave us the wonderful opportunity to work on this wonderful project that also allowed us to conduct a great deal of research. We also wish to express our gratitude to the computer science and engineering professors whose suggestions enabled us to move past a lot of seemingly insurmountable obstacles.

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.918 ⁜ Vol. 12, Issue 2, February 2023

DOI: 10.17148/IJARCCE.2023.12231

REFERENCES

[1] S. Gollapudi, "Practial Machine Learning," Packt Publishing Ltd., 2016.
[2] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning.(Report),” Nature,vol. 521, no. 7553, pp. 436, May 2015, 2015.
[3] J. Shah, "Neural Networks for Beginners: Popular Types and Applications," https://blog.statsbot.co/neural-networks-for-beginnersd99f2235efca,[January 12, 2019, 2017].
[4] O. Mogren, “C-RNN-GAN: Continuous recurrent neural networks with adversarial training,” CoRR, vol. abs/1611.09904, 2016.
[5] Y. Chen, “Towards Explaining Neural Networks,” Faculty of Science,Utrecht University, Utrecht, Netherlands, 2017.
[6] Tomas Mikolov, Martin Karafiat, Lukas Burget et al., “Recurrent neuralnetwork based language model,” INTERSPEECH, vol. 2, pp. 4, 2010.
[7] Y. M. G. Costa, L. S. Oliveria, and C. N. S. Jr., "An Evaluation ofConvolutional Neural Networks for Music Classification UsingSpectrograms," 52,https://www.sciencedirect.com/science/article/pii/S1568494616306421
[8] S. Dieleman, P. Brakel and B. Schrauwen, "Audio-based music classification with a pretrained convolutional network", in 12th International Society for Music Information Retrieval Conference,ISMIR, 2011, pp. 669-674.
[9] T.L. Li, A.B. Chan and A. Chun, "Automatic musical pattern feature extraction using convolutional neural network", in Proceedings of the International MultiConference of Engineers and Computer Scientists, Hong Kong, 2010.
[10] H. Lee, Y. Largman, P. Pham and A. Ng, "Unsupervised feature learning for audio classification using convolutional deep belief networks", in *Advances in neural information processing systems*, NIPS, 2009, pp. 1096-1104.