



International Journal for Innovative Engineering and Management Research

A Peer Reviewed Open Access International Journal

www.ijiemr.org

COPY RIGHT



ELSEVIER
SSRN

2019IJIEMR. Personal use of this material is permitted. Permission from IJIEMR must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. No Reprint should be done to this paper, all copy right is authenticated to Paper Authors

IJIEMR Transactions, online available on 1st Jun 2019. Link

[:http://www.ijiemr.org/downloads.php?vol=Volume-08&issue=ISSUE-06](http://www.ijiemr.org/downloads.php?vol=Volume-08&issue=ISSUE-06)

Title: **FACIAL EXPRESSION BASED MUSIC RECOMMENDATION SYSTEM USING CNN**

Volume 08, Issue 06, Pages: 23–26.

Paper Authors

Manjula K, Diya Fathima M, Harshitha S.

Sri Krishna Institute of Technology, Bangalore.



USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per **UGC Guidelines** We Are Providing A Electronic Bar Code

FACIAL EXPRESSION BASED MUSIC RECOMMENDATION SYSTEM USING CNN

¹Manjula K., ²Diya Fathima M., ³Harshitha S.

¹Assistant Professor, Department of ISE, Sri Krishna Institute of Technology, Bangalore.

^{2,3}Student, Department of ISE, Sri Krishna Institute of Technology, Bangalore.

ABSTRACT

These days development of Music Recommendation system are being evolved based on various parameters such as Audio-based, Lyrics-based etc. Music Recommendation system helps the music listeners for faster access of their preferred music. However, the previous methods used different Machine learning algorithms to access Music based on different parameters. Therefore, we have introduced a concept of recommending music based on the Facial expressions that the user presents such as, happy, sad, angry, etc. For the proposed system, the dataset can be taken as a set of images for each class which should be trained. In order to implement this system, we propose an algorithm which is a deep learning algorithm known as CNN (Convolutional Neural Network).

KEYWORDS

Music Recommendation, Facial Expressions, Machine Learning, Deep Learning, CNN.

1. INTRODUCTION

Listening to Online music has taken the digitization to another level. Music listeners tend to listen to music based on recommendations as per their desires. The music listeners can access their desired music using technologies like Big Data, Cloud, etc. In the previous years, numerous approaches have emerged for the conviction of Music recommendation. The furtherance in the aspect of cloud computing have assisted the music listeners to access unbounded variety of music.

Since there is difficulty in playing music only based on unsystematic basis, users' judge on account of accessibility. Hence, we instigate the utilization of music in the form of capturing the Facial expressions. In this paper, we mainly concentrate on the Facial Expressions of the user such that different expressions are captured in order to recommend the most appropriate music reliable for the user. Songs are collected and trained from MIDI (Music Instrument Digital Interface) or MP3 files or Million song dataset.

However, preceding experiments have shown their conclusions in the aspects like identifying music notes, by gathering the information about the music, etc. As mentioned earlier, we are proposing an algorithm based on Deep Learning called as CNN (Convolutional Neural Network) which is nothing but a class of deep neural network, most customarily pertained for visual imagery. CNN take a different approach towards regularization. They take ascendancy of the hierarchical pattern in data and congregate more complex patterns using compact and effortless patterns. CNN use adequately pre-processing compared to other image classification algorithms. This means that the network assimilate the filters that in traditional algorithms were hand-engineered. CNN involves three major methods through which selection and classification of images are performed. Those include,

- 1) Filtering of the images in the dataset.
- 2) ReLU (Rectified Linear Unit) activation function.
- 3) Pooling which is used to control the overfitting.

By using the above mentioned CNN methods and techniques, we can obtain the well-versed version of music recommendation system using facial expressions.

2. RELATED WORK

A. What to play next? RNN-based Music Recommendation System

In this paper, music would be recommended based on audio-based and lyrics-based. For the audio-based recommendation, we use the RNN-LSTM algorithm which is mainly used to retain the previously output for long-term dependencies. In [1], an improved algorithm is used such that it estimates the equivalences of various songs. The recommendations are made by perceiving the contents of the song. Since this paper uses Million song Dataset, management of such huge data is highly burdensome.

B. Advances in Neural Information Processing Systems

This paper shows how the information passes and processes through the neural networks. In the neural networks, there are multiple layers in the hidden network for computation of a chunk of data. In [2], since at a particular time interval, a huge amount of data cannot be processed, it becomes a hindrance in computing data which possess a large set of information which increases the complexity of the system evaluating neural networks.

C. Human Facial Expression Detection from Detected in Captured Image Using Back-Propagation Neural Network

In this paper, the facial expressions are detected using an approach which is grounded on an add-boosted classifier for the detection of facial expressions. Back propagation neural network is been used for identification and matching of tokens. In [3], the disadvantage is that the generation of the tokens based on the best threshold value is quite a tedious job.

3. PROPOSED METHOD

In our paper, we have explained the methods and approaches that will take place in order to obtain the results for the music recommendation system based on the facial expressions. In deep learning, there are multiple layers where each level is formulated such that it is able to reconstruct its input data into a slightly digestible representation. In our system, the raw image is represented in a matrix of some particular pixels. As we are considering an improved Deep Neural Network algorithm which is nothing but an Artificial Neural Network (ANN) with ample number of layers linking the input and output layers. Here, we use three important layers namely the input layer, the hidden layer and the output layer. Each layers contain certain functionalities to perform which will help provide the required output. It moves from one layer to the other layer only after scheming the probability of the previous layers.

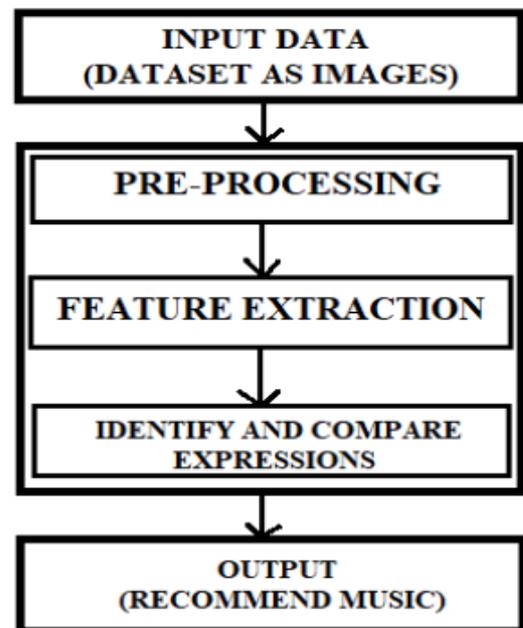


Figure 1. Procedural Methods of Facial Expression

The above Figure 1 gives a brief description about the various methods that are handled in each layer. Firstly, the images should be trained which possesses images with different facial expressions such as

happy, sad, angry, etc. In the input layer, we capture a random image of a person with some particular facial expression. Followed by, this captured image is moved to the hidden layer. In the hidden layer, there are three internal operations taking place, namely Pre-processing, Feature Extraction, identification and comparison of expressions with the trained dataset. In pre-processing, the unwanted part of the image is eliminated, i.e., the backdrop part of the image. In the feature extraction stage, all the features which will be necessary for the further analysis such as edges of the mouth, eyebrows, etc are extracted. Finally in the Identification and Comparison stage, the extracted features are identified and compared with the previously trained dataset. The image that suits with the trained dataset and provides the highest accuracy is chosen and according to the matching expression, the corresponding music will be recommended and the music will be played.

A. Use Case Diagram of the Proposed System:

The Figure 2 represents the Use Case diagram of the proposed system which gives the brief description about the interactions taking place between the User and the system.

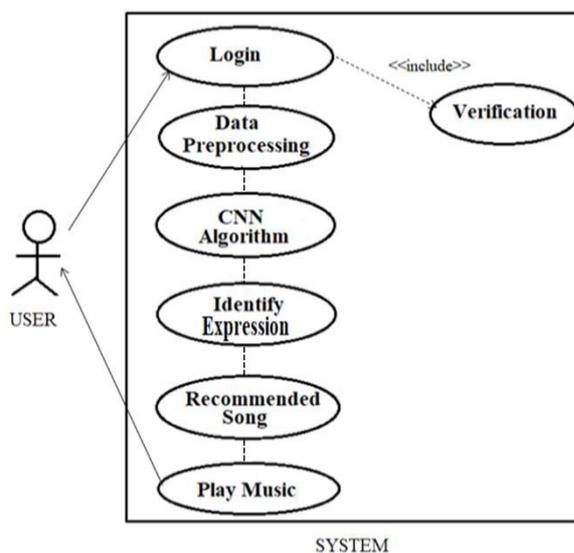


Figure 2. Use Case Diagram of the Proposed System

B. Algorithm

Step 1: The image is fragmented into 3 color-channels RGB (Red, Green and Blue). All the three color-channels are delineated separately to some particular image pixels.

Step 2: Every convolution filter constitutes a variant characteristic and the Convolutional Neural Network algorithm masters which characteristics incorporates the deriving reference.

Step 3: ReLU is an activation function. The foremost objective of ReLU is to detach all the negative values from the convolution. All the positive values which are extracted prevail unchanged but all the negative values will be converted to zero.

Step 4: In the pooling layer, we compress the image stack into a compact size.

Step 5: Predictions are based on the output data which are obtained by collating the acquired values with the dataset.

C. System Design

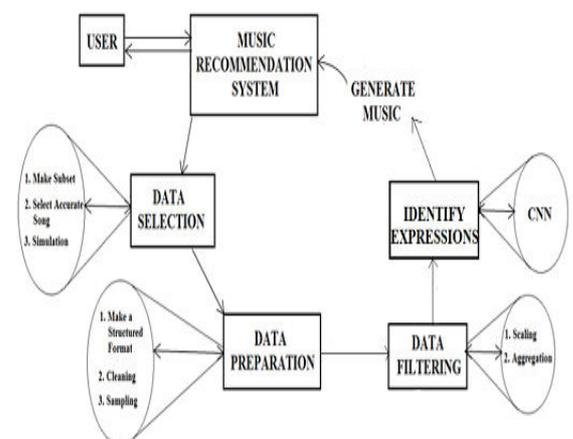


Figure 3. Design of the Proposed System

The Figure 3. shows the diagrammatic representation of the internal working of the proposed system. The user can register to the

music recommendation system by submitting the required user credentials with a valid password. Once the user is registered, the user can login by entering the valid username and password which was submitted at the time of registration. After logging in to the music recommendation system, it captures the user's image for a particular time interval. Then the image undergoes four major stages which are as follows,

1. Data Selection
2. Data Preparation
3. Data Filtering
4. Identify Expressions

In Data Selection, three phases takes place which are as follows

1. Make a subset- Here, the image is divided into chunks and subset is made of this data and successive frames of variant sizes will be produced.
2. Select Accurate Image- Now, the most accurate part of the image which is needed for further analysis is considered.
3. Simulation- Comparison of the captured image with the trained dataset.

In Data Preparation, three phases takes place which are as follows

1. Make a structured format- Compress the image such that only the required part of the image is considered.
2. Cleaning- Undesirable part of the captured image is removed by using a Least Mean Square method.
3. Sampling- Producing and combining of data from the subset.

In Data Filtering, two phases takes place which are as follows

1. Scaling- Selecting the right feature.
2. Aggregation- Collate or combine the right feature.

In the fourth phase, we identify the expressions extracted from the captured image and we use the CNN algorithm for

comparison and recommendation of music.

4. CONCLUSION

The music recommendation system based on the Facial Expression of the user using Convolutional Neural Network is proposed in this paper. Here the expressions are captured and are subjected to various phases so that a clear expression is obtained and are compared with the images that are trained in the dataset. After the confirmation of the facial expression, whether it is Happy, Sad, Angry, etc. the music will be recommended and played accordingly.

5. REFERENCE

- [1] Miao Jiang, Ziyi Yang and Chen Zhao, *What to play next? A RNN-based music recommendation system*, 978-1-5386-1823-3/17, 2017.
- [2] A Van den Oord, S Dieleman, B Schrauwen, "Advances in Neural Information Processing Systems", 2643-2651, 2013
- [3] Jagdish Lal Raheja and Umesh Kumar, "Human Facial Expression Detection from Detected in Captured Image Using Back-Propagation Neural Network", *International Journal of Computer Science and Information Technology (IJCSIT)*, Vol. 2, No.1, February 2010.
- [4] Raman Maini and J.S.Sohal, "Performance evaluation of Prewitt Edge detector for noisy image", *Journal on Graphics, Vision and Image Processing*, Vol 6, December-2006.
- [5] Thierry Bertin-Mahieux, Daniel P.W.Ellis, Brian Whitman and Paul Lamere, "The Million Song Dataset", *Proceedings of the 12th International Society for Music Information Retrieval Conference*, 2011.



International Journal for Innovative Engineering and Management Research

A Peer Reviewed Open Access International Journal

www.ijiemr.org