MANTECH Publications

Journal of Software Engineering & Software Testing
Volume 4 Issue 1
ISSN: 2457-0516 (Online)

# *Emotion Based Music Player*

***Vinayak Bali, Shubham Haval*, Snehal Patil, R. Priyambiga***
*Sanjay Ghodawat University (India)*
***Corresponding author's email id:*** *shubhamhaval619@gmail.com**
*DOI: http://doi.org/10.5281/zenodo.2647696*

***Abstract***

*Listening to music affects the human brain activities. Emotion based music player with automated playlist can help users to maintain a particular emotional state. This research proposes an emotion based music player that creates a playlists based on captured photos of the user. Manual sorting of a playlist and annotation of songs, in accordance with the current emotion, is more time consuming and quite tedious. Numerous algorithms have been implemented to automate this process. However, existing algorithms are slow, increase cost of the system by using additional hardware and have quite very less accuracy. This paper presents an algorithm that not only automates the process of generating an audio playlist, but also to classify those songs which are newly added and the main task is to capture current mood of person and to play song accordingly. This enhances the system's efficiency, faster and automatic. The main goal is to reduce the overall computational time and the cost of the designed system. It also aims at increasing the accuracy of the system. The most important goal is to make change the mood of person if it is a negative one such as sad, depressed. This model is validated by testing the system against user dependent and user independent dataset.*

***Keywords:*** *Convolution neural network, Long Short term memory, Emotion detection, audio classification, hidden layers, Max-pooling.*

## INTRODUCTION

Expressing and recognizing emotions of human are very much important in communication system [3]. Human beings have the ability to express and recognize emotions. Computer seeks to identify the human emotions either by image analysis or through sensors [6]. In our day to day

life and in our professional life we interact with many people face to face or indirectly by phone calls, sometimes it is necessary for people to be aware of their present emotions of the person with whom they are interacting. Human emotions are classified as: surprise, fear, anger, happy, sad, disgust and neutral [3].

Facial movement [1] and the tone of speech play a major role in expressing emotions. The physique and tone of the face tells the energy in the utterance of speech, which can be firstly modified to communicate different feelings. Humans can easily recognize these changes in signals along with the information felt by any other sensory organs. This project analyses the use of image or sensors or speech to capture the emotions.

Music plays [4] a vital role in enhancing an individual's life as it is an important medium of entertainment for music lovers and listeners [15] and sometimes even imparts a therapeutic approach. "Where words fail music speaks", and hence it can change person's negative emotion simultaneously and slowly into a positive mood [7].

Emotions can be expressed through gestures, speech, facial expressions, body language etc. For the system to understand the user's mood, we use facial expression [3]. Using the mobile device's camera, we can capture the user's facial expression. There are many emotion recognition systems which take captured image as input and determine the emotion. For this application, we are using neural networks for recognition of emotion[9][8].

## IMPLEMENTATION

The proposed algorithm revolves automated music recommendation system that plays song according to the mood or current emotion of the person. The person's photo is captured whenever the application gets open; hence current emotion is captured and detected. According to the information given by the image, the song is played related to the emotion.

The songs whichever present in the phone are already classified into 7 different classes [7] such as happy, sad, anger, surprise, fear, disgust and neutral. The newly added songs are also classified dynamically to appropriate mood. It is composed of three modules: Facial expression recognition module, Song emotion recognition module and System integration module. Facial expression recognition and audio emotion recognition

modules are two mutually exclusive modules. Hence, system integration module maps two modules to find the correct match of detected emotion. ***(See Figure:-1)***

***A. Data Set***

The Raw dataset is downloaded one by one from Google images for seven emotions. Extra dataset is taken from Kaggle datasets for facial expression detection.

***B. Trained Dataset***

Before processing the model, the training and testing phases is undergone. Trained dataset are those which are taught to the model or which learn

At the time of training, system takes dataset of faces (images) with their respective expression; eye should be in centre location mostly and learns a set of weights, which splits the facial expressions for classification. ***(See Figure:-2)***

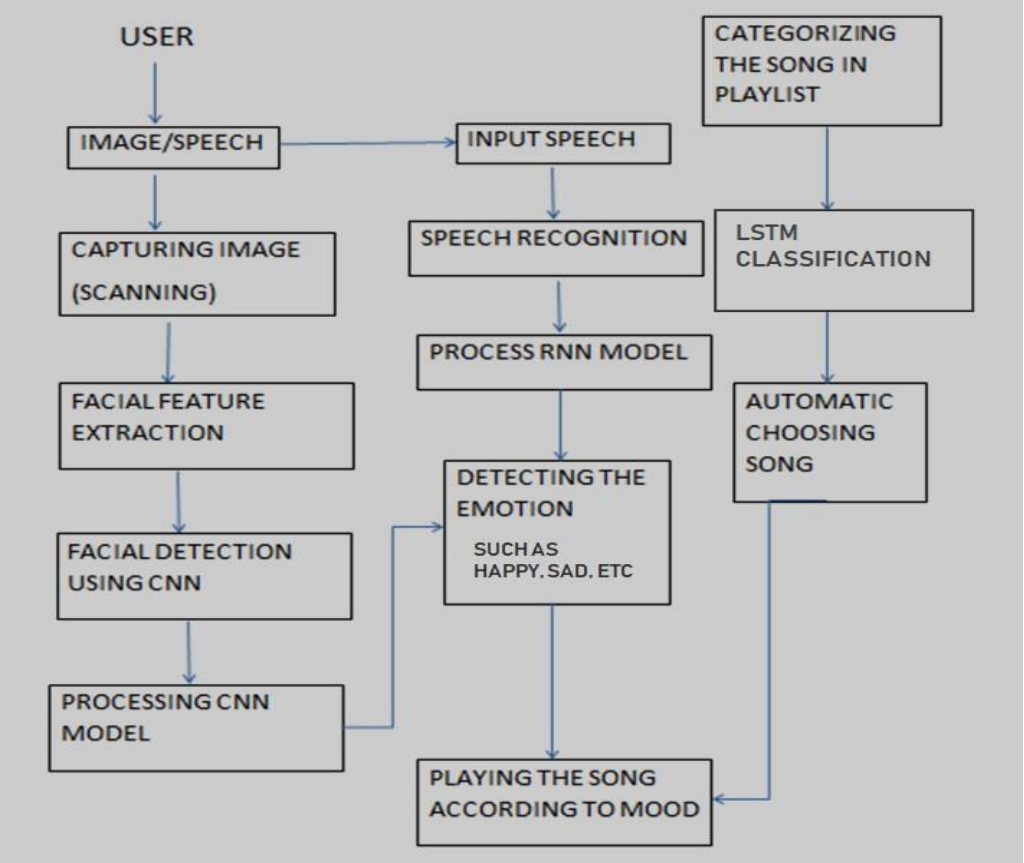


***Figure 1: Block Diagram***

***Sample trained data set is:***


***Figure 2: Trained dataset***

***For training, the sequence is:***

1. Spatial normalization
2. Synthetic samples generation
3. Image cropping
4. Down-sampling
5. Intensity normalization.

### *C. Test Data*

At the time of testing, classifier takes images of face with respective eye center locations, and it gives output as predicted expression by using the weights learned during training.


***Figure 3: Test dataset***

***For recognizing an unknown image (testing), the sequence is:***

1. Spatial normalization
2. Image cropping
3. Down-sampling
4. Intensity normalization

### *D. Convolution Neural Network*

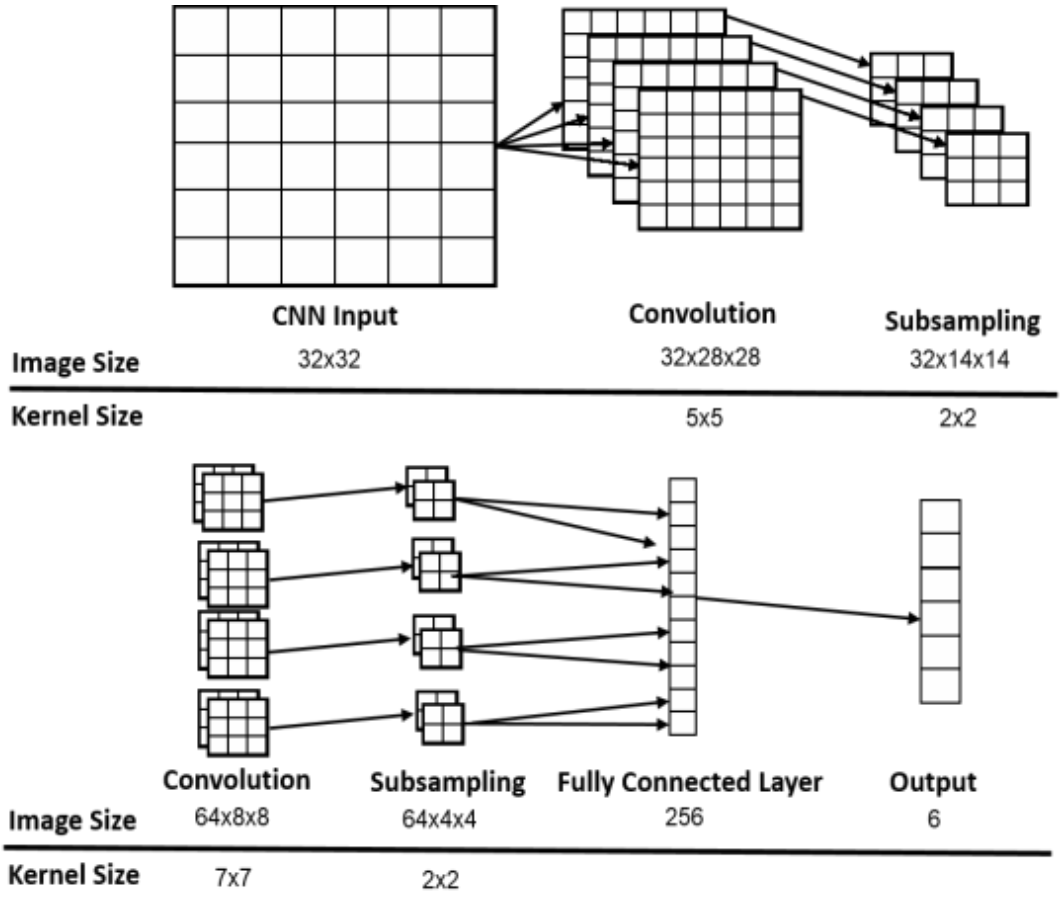


***Figure 4: Architecture of proposed Convolution Neutral Network. It has five layers: first layer is convolution, second layer is sub-sampling, the third layer is convolution, fourth layer is sub-sampling, fifth layer is fully connected layer and final responsible for classifying facial image.***

**FINAL RESULTANT MODEL**

All the photos present in dataset are firstly converted to grayscale, for making preprocessing and detection more efficiently, faster and easier. Each input image is in form of pixels (e.g. 48x48). Now the pixel represented images are sent to the convolution layers (hidden layers). In Between each layer maximum pooling is done, the purpose of doing so, is to down-sample the input data or image, reducing the dimensions and allows assumption to be made about features contained in sub regions. This is done to avoid over-fitting. As well as it reduces computational cost by reducing number of parameters to learn. Example, if input image is of matrix 4x4 representation and let's say output we want is in 2x2, then pooling is performed in between all hidden layers. After that data is sent to dense layer, to prevent over-fitting. Dropout technique is used to reduce over-fitting in neural networks. The output layer conveys the detected class. Let's say if the detected expression is happy, then the next step is to select anyone training dataset for music model. Now, the dataset is trained according to the match for playing music. LSTM neural network is used for classifying the songs. One hot encoding is performed to represent categorical variables into binary vectors, so as to make the classification faster and better. Then the song is played according to the current mood of the person.

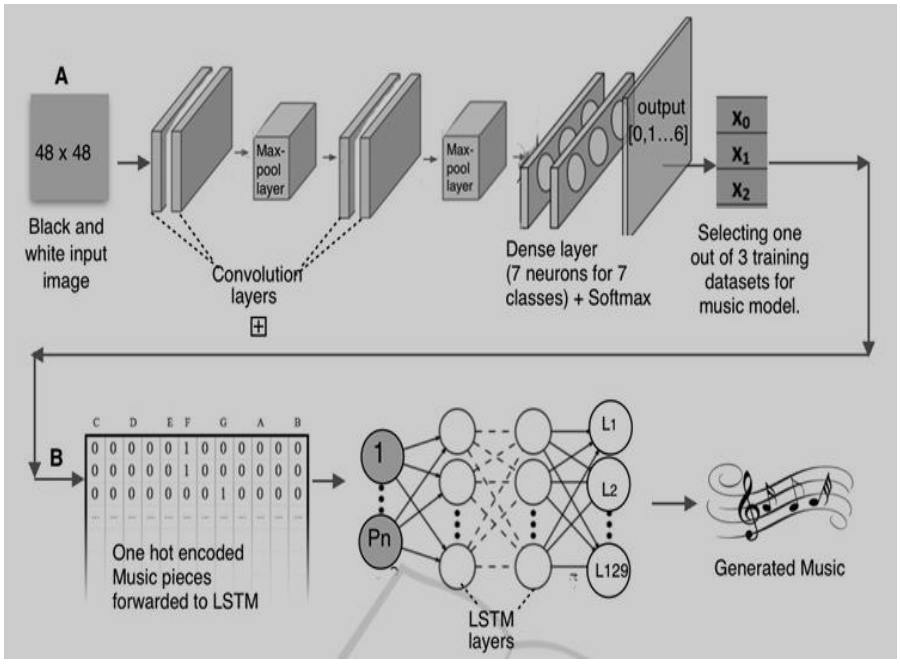


***Figure 5: Final Model (Resultant)***

## RESULT ANALYSIS


***Figure 6: Welcome Activity***


***Figure 7: Option Activity***


***Figure 8: Playlist***

***Figure 9: Capturing image***

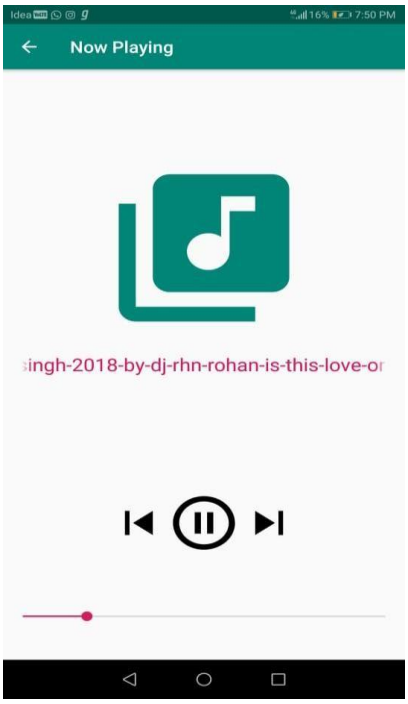


***Figure 10: Song played according to mood (OUTPUT)***

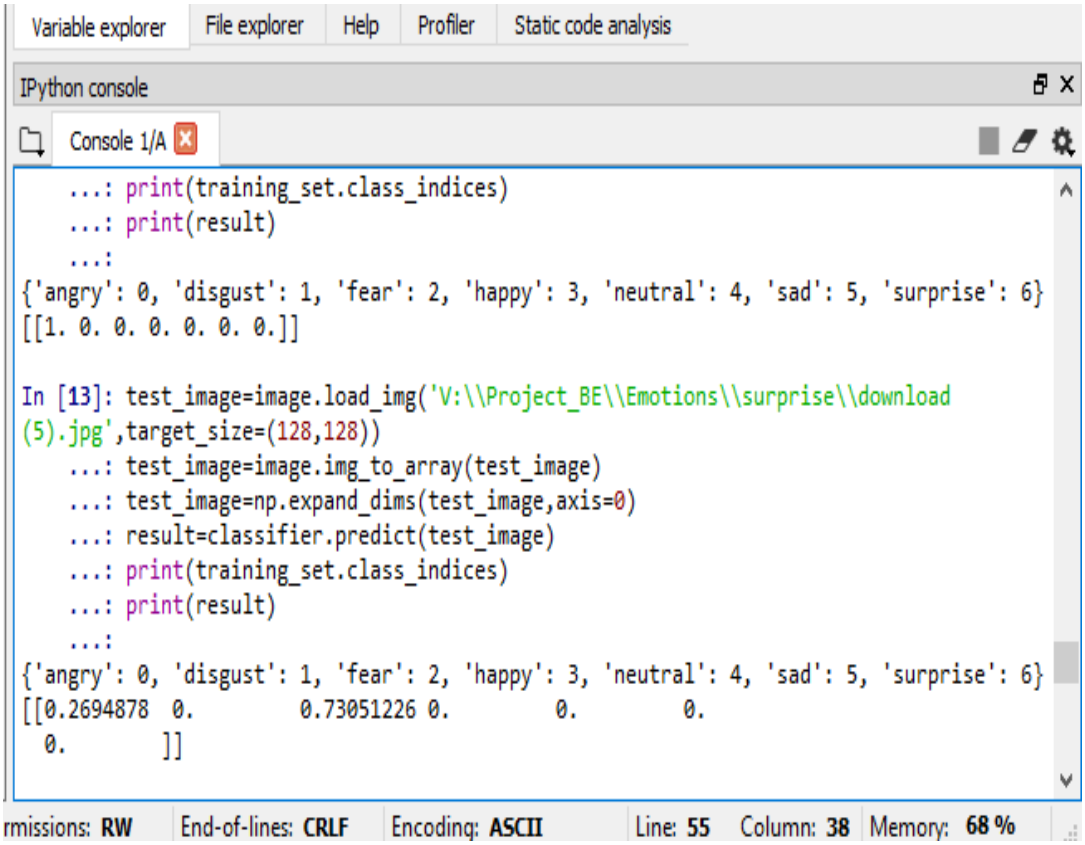


***Figure 11: Backend (emotion detection)***

## CONCLUSION

The Emotion Based Music Player is used to automate and give a best music player experience for end user. Application solves all the basic needs of music listeners without troubling them as existing applications do. It uses technology to increase the interaction of the system with the user in numerous ways. It eases the work of user by capturing the image using phone's camera, detecting their emotion and suggesting a customized playlist with advanced features. The user's negative or bad thoughts are slowly converted to positive thoughts by changing the song from low tone to excited tone.

## FUTURE SCOPE

In future Music Player can be enhanced with Google play music, so songs which are not present in local storage can also be played and to access the whole application in speech based. The Emotion Based Music System will be of great advantage to users looking for music based on their mood and emotional behavior. It will help reduce the searching time for music thereby reducing unnecessary time and hence increasing the overall accuracy and efficiency of the system. The system will not only reduce physical stress but will also act as a boon for the music therapy systems and may also assist the music therapist to treat the patient. In future it can also be used to detect the sleepy mood of the driver, driving the car and many more uses. Also with its additional features mentioned above, it will be a complete system for music lovers and listeners.

## REFERENCES

I. Y. Wu, H Liu, and H.Zha,―Modeling facial expression space for recognition,‖ in 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2005. (IROS 2005), 2005, pp. 1968–1973.

II. S. Z. Li and A. K. Jain, Handbook of Face Recognition. Springer Science & Business Media, 2011.

III. Mireille Besson, Frederique Faita, Isabelle Peretz, A- M Bonnel, and Jean Requin. Singing in the brain: Independence of lyrics and tunes. Psychological Sci- ence, 9(6):494–498, 1998.

IV. S. Dornbush, K. Fisher, K. McKay, A. Prikhodko and Z. Segall, Xpod- A Human Activity and Emotion Aware Mobile Music Player, UMBC Ebiquity, November 2005.

V. G. Heamalathal, C.P. Sumathi, A Study of Techniques for Facial Detection and Expression Classification, International Journal of Computer Science & Engineering Survey(IJCSES) Vol.5, No. 2, April 2014.

VI. Alvin I. Goldmana, b.Chandra and SekharSripadab, ―Simulationist models of face-based emotion recognition.

VII. Thayer, The biopsychology of mood & arousal, Oxford University Press ,1989. S. Demyanov, J. Bailey, R. Kotagiri, and C. Leckie, ―Invariant back-propagation: how to train a transformationinvariant neural network,‖ arXiv:1502.04434 [cs, stat], 2015.

VIII. D. C. Cirean, U. Meier, J. Masci, L. M. Gambardella, and J. Schmid-huber, ―Flexible, high performance convolutional neural networks for image classification,‖ in Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume Two, ser. IJCAI'11. AAAI Press, 2011, pp. 1237–1242.

IX. Colah's blog, Understanding LSTM Networks Posted on August 27, 2015.

X. Debaditya et al., 2015] Debaditya, R., Sri, R. M. K., and Krishna, M. C. (2015). Feature selection using deep neural networks.

XI. Yadan O, Adams K, Taigman Y, et al. Multi-GPU Training of ConvNets[J]. Eprint Arxiv, 2013.

XII. J.E. Moody. The effective number of parameters: An analysis of generalization and regularization in nonlinear learning systems. In Advances in Neural Information Processing Systems, volume 4, pages 847-854. Morgan Kaufmann, 1992.

XIII. Yan, P., Huang, R.: Artificial Neural Network — Model, Analysis and Application. Anhui Educational Publishing House, Hefei."Impact of Music, Music Lyrics, and Music Videos on Children and Youth". Pediatrics.

124 (5): 1488–1494. 19 October 2009.

XIV. Radford, C. (1989). "Emotions and music: A reply to the cognitivists". The Journal of Aesthestics and Art Criticism. Wigram T. Indications in music therapy: Evidence-based practice. British Journal of Music Therapy 2002.

XV. Hinton, G. E.; Srivastava, N.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R.R. (2012). "Improving neural networks by preventing coadaptation of feature detectors". Startups, Requests for (2016-0812). "Deep Learning in Healthcare: Challenges and Opportunities". Medium. Retrieved 2018-04-10.

XVI. Bengio, Y. (1991). "Artificial Neural Networks and their Application to Speech/Sequence Recognition". McGill University Ph.D. thesis. Abdel-Hamid, O.; et al. (2014)."Convolutional Neural Networks forSpeechRecognition" (PDF). IEEE/ACM Transactions on Audio, Speech, and Language Processing.

XVII. B. Fasel, ―Robust face analysis using convolutional neural networks,‖ in 16th International Conference on Pattern Recognition, 2002. Proceed- ings, vol. 2, 2002, pp. 40–43 vol.2.

XVIII. F. Beat, ―Head-pose invariant facial expression recognition using con- volutional neural networks,‖ in Fourth IEEE International Conference on Multimodal Interfaces, 2002. Proceedings, 2002, pp. 529–534.

XIX. M. Matsugu, K. Mori, Y. Mitari, and Y. Kaneda, ―Subject independent facial expression recognition with robust face detection using a convo- lutional neural network,‖ Neural Networks: The Official Journal of the International Neural Network Society, vol. 16, no. 5, pp. 555–559, 2003.

XX. P. Zhao-yi, W. Zhi-qiang, and Z. Yu, ―Application of mean shift algorithm in real-time facial expression recognition,‖ in International Symposium on Computer Network and

Multimedia Technology, 2009. CNMT 2009, 2009, pp. 1–4.

XXI. P. Lucey, J. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews,―The extended cohn-kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression,‖ in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Work-shops (CVPRW), 2010, pp. 94–101.

XXII. C.-D. Caleanu, ―Face expression recognition: A brief overview of the last decade,‖ in 2013 IEEE 8th International Symposium on Applied Computational Intelligence and Informatics (SACI), 2013, pp. 157–161.

**Cite this Article**

Vinayak Bali, Shubham Haval, Snehal Patil, R. Priyambiga (2019). **Emotion Based Music Player** Journal of Software Engineering & Software Testing, 4(1), ISSN: 2457-0516 (Online) 18- 30

*http://doi.org/10.5281/zenodo.2647696*

## AUTHORS' PROFILE

***[1] Vinayak Bali,*** *Student*
*Department: Computer Science and Engineering*
*College: Sanjay Ghodawat University (India)*

***[2] Shubham Haval,*** *Student*
*Department: Computer Science and Engineering*
*College: Sanjay Ghodawat University (India)*
*Email id: shubhamhaval619@gmail.com*

***[3] Snehal Patil,*** *Student*
*Department: Computer Science and Engineering*
*College: Sanjay Ghodawat University (India)*

***[4] R.Priyambiga,*** *Assistant Professor*
*Department: Computer Science and Engineering*
*College: Sanjay Ghodawat University (India)*