iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** XII **Month of publication:** December 2018

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue XII, Dec 2018- Available at www.ijraset.com

Emotion Based Music Player

Bharti Rathi[1], Girija Joshi[2], Tejas Kamble[3]
[1, 2, 3]*Computer Department, Savitribai Phule Pune University, India*

Abstract: Machine learning is a field of computer science that uses statistical techniques to give computer systems the ability to "learn" (e.g., progressively improve performance on a specific task) with data, without being explicitly programmed.
Songs, as a medium of expression, have always been a popular choice to depict and understand human emotions. Speech is the most ancient and natural way of expressing feelings, emotions and mood and its processing requires high computation, time and cost. Emotion based music player application is based on real-time extraction of facial expressions as well as extracting audio features from songs to classify into a specific emotion that will generate a playlist automatically.
Emotion based music player application contains three modules: Emotion Module, Music Classification Module and Recommendation Module. The Emotion Module takes an image of the user's face as an input and makes use of deep learning algorithm to identify their mood. The Music Recommendation Module suggest songs to the user by mapping their emotions to the mood type of the song, taking into consideration the preferences of the user. The proposed application has an accurate and efficient model that will generate playlist based on current emotional state and behavior of the user.
Keywords: Recommendation systems, Emotion recognition, Music information retrieval, Convolutional neural networks, Multi-layer neural network.

I. INTRODUCTION

In this paper, we focus on automatically annotating faces in images based on the ambiguous supervision from the associated captions gives. Some pre-processing steps need to be conducted before performing face naming.

Specifically, faces in the images are automatically detected using face detectors and names in the captions are automatically extracted using a name entity detector. Here, the list of names appearing in a caption is denoted as the candidate name set. Even after successfully performing these pre-processing steps, automatic face naming is still a challenging task.

The faces from the same subject may have different appearances because of the variations in poses, illuminations, and expressions. Moreover, the candidate name set may be noisy and incomplete, so a name may be mentioned in the caption, but the corresponding face may not appear in the image, and the correct name for a face in the image may not appear in the corresponding caption.

Each detected face (including falsely detected ones) in an image can only be annotated using one of the names in the candidate name set or as null, which indicates that the ground-truth name does not appear in the caption. and also we have to develop emotion musical system we have detect emotions using different sounds and display the list of sounds to that particular emotions

II. LITERATURE SURVEY

This paper presents a technique for automatically detecting human faces in digital color images. This is two-step process which first detects regions contain human skin in the color image and then extracts information from these regions which might indicate the location of a face in the image. The skin detection is performed using a skin filter which relies on color and texture information. The face detection is performed on a grayscale image containing only the detected skin areas. A combination of thresh Holding and mathematical morphology are used to extract object features that would indicate the presence of a face. The face detection process works predictably and fairly reliably, as test results show.

Each detected face (including falsely detected ones) in an image can only be annotated using one of the names in the candidate name set or as null, which indicates that the ground-truth name does not appear in the caption. and also we have to develop emotion musical system we have detect emotions using different sounds and display the list of sounds to that particular emotions.

This proposed system based on extraction of facial

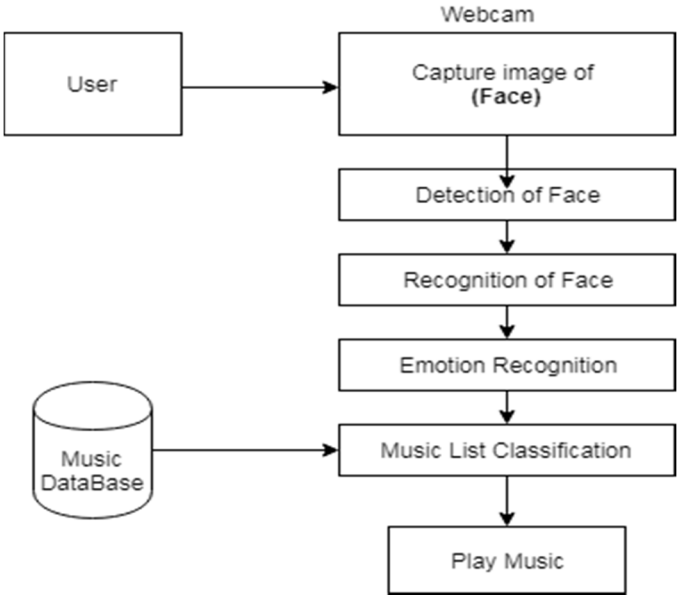
Expressions as well as extracting audio features from songs to classify into a specific emotion that will generate a playlist automatically.

In this paper, present an affective cross-platform music player, EMP (Electro-magnetic Pulse), Which recommends music based on real-time mood of the user.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue XII, Dec 2018- Available at www.ijraset.com

III. PROPOSED SYSTEM

The purpose of face naming is to surmise the right emotions for every face. In this we propose two new strategies to adequately take care of this issue by taking in two discriminative proclivity lattices from these feebly named pictures.
In this we propose a new scheme for automatic sound detection with caption-based supervision.



General flow of proposed Emotion Based Music Player

1) Webcam capture image of user (face) by using camera.
2) After capturing image it detect that image in database.
3) Then in between the capturing image and storing images in the database mapping takes place by matching their pixels.
4) Emotion get recognized by matching pixels.
5) Database fetch the playlist of songs which is available in the database as per the emotion detected and music get played.

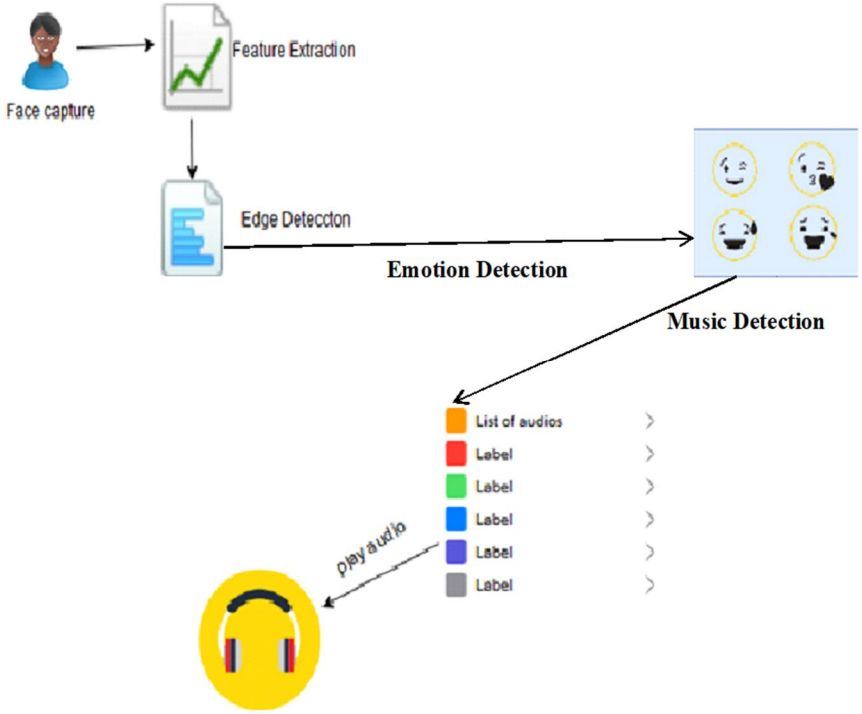


Fig. Architecture of system

A. Emotion module

The Emotion Module takes an image of the user's face as an input and makes use of deep learning algorithms to identify their mood.

B. Music Classification module

The Music Classification Module makes use of audio features.

C. Recommendation module

The Recommendation Module suggests songs to the user by mapping their emotions to the mood type of the song, taking into consideration the preferences of the user using 3 convolutional layer in neural network.

This application is based on real-time execution of facial expressions as well as extracting audio features from songs to classify into a specific emotion that will generate a playlist automatically.

IV. FUTURE SCOPE

In the future, we plan to enhance the system with some other estimation tools and statistical analysis. This might be used not only by public users but also by decision makers in the local municipalities. Moreover, since the system is developed with open standards and open sources, it is easily extended with future technologies according to users' needs.

V. CONCLUSION

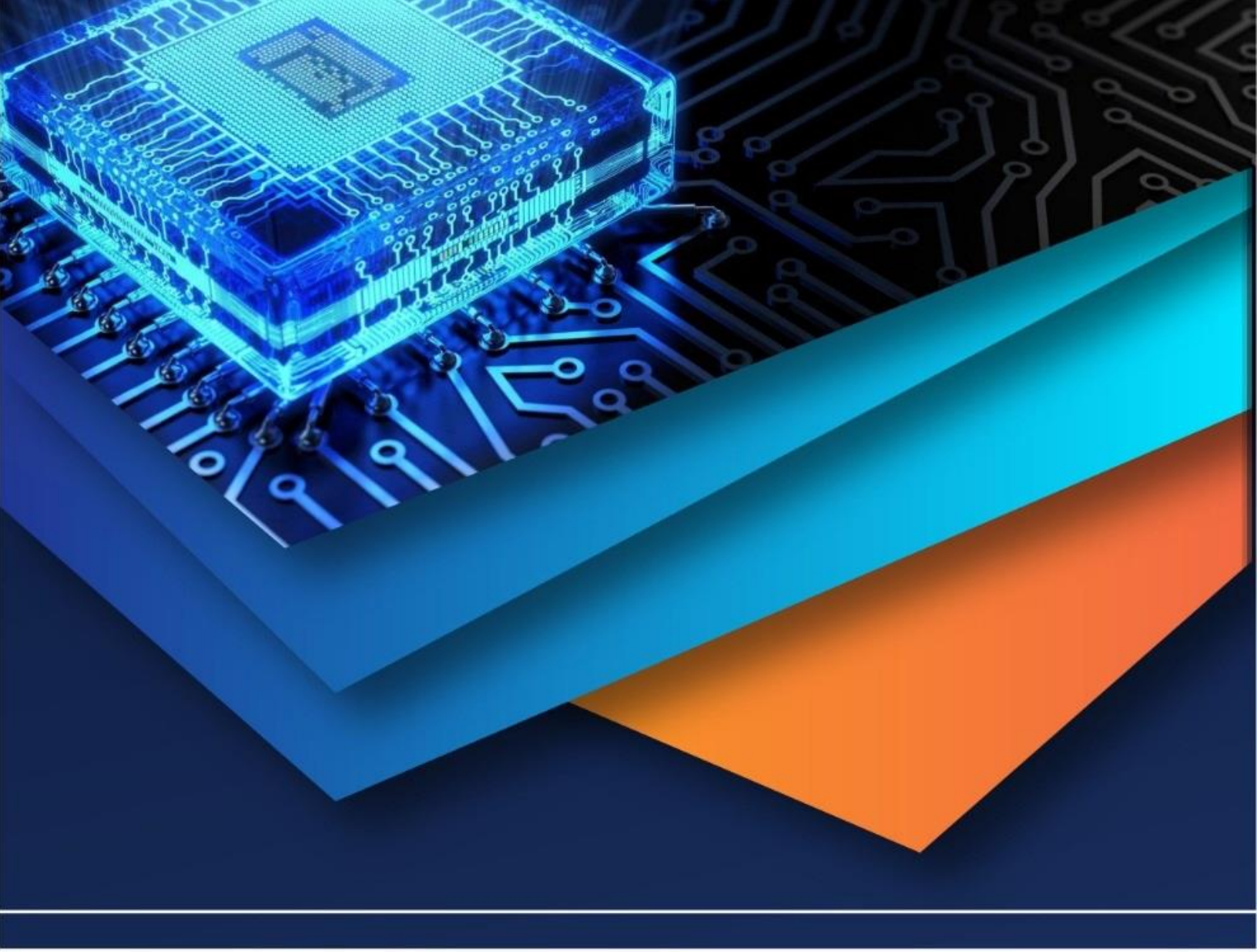
The Automatic face and sound Detection. This method is secure, reliable and easy to use. Automated Face and sound Systems based on face recognition techniques thus proved to be time saving and secured.

This system can also be used to identify a different emotions. An automatic face and sound management system is a necessary tool for any organization. Most of the existing systems are time consuming and require for a semi manual work. This approach aims to solve the issues by integrating face recognition in the process.

REFERENCES

[1] P. Viola and M. J. Jones, "Robust real-time face detection," Int. J.Comput.Vis., vol. 57, no. 2, pp. 137–154, 2004.
[2] G. Liu, Z. Lin, and Y. Yu, "Robust subspace segmentation by low-rankrepresentation," in Proc. 27th Int. Conf. Mach. Learn., Haifa, Israel,Jun. 2010, pp. 663–670.
[3] T. L. Berg et al., "Names and faces in the news," in Proc. 17th IEEEComput. Soc. Conf. Comput. Vis. Pattern Recognit., Washington, DC,USA, Jun./Jul. 2004, pp. II-848–II-854.
[4] D. Ozkan and P. Duygulu, "A graph based approach for naming facesin news photos," in Proc. 19th IEEE Comput. Soc. Conf. Comput. Vis.Pattern Recognit., New York, NY, USA, Jun. 2006, pp. 1477–1482.
[5] P. T. Pham, M. Moens, and T. Tuytelaars, "Cross-media alignment ofnames and faces," IEEE Trans. Multimedia, vol. 12, no. 1, pp. 13–27,Jan. 2010.
[6] M. Guillaumin, J. Verbeek, and C. Schmid, "Multiple instance metriclearning from automatically labeled bags of faces," in Proc. 11th Eur.Conf. Comput. Vis., Heraklion, Crete, Sep. 2010, pp. 634–647.
[7] J. Luo and F. Orabona, "Learning from candidate labeling sets," inProc. 23rd Annu. Conf. Adv. Neural Inf. Process. Syst., Vancouver, BC,Canada, Dec. 2010, pp. 1504–1512.
[8] X. Zhang, L. Zhang, X.-J.Wang, and H.-Y. Shum, "Finding celebritiesin billions of web images," IEEE Trans. Multimedia, vol. 14, no. 4,pp. 995–1007, Aug. 2012.
[9] Z. Zenget al., "Learning by associating ambiguously labeled images,"in Proc. 26th IEEE Conf. Comput. Vis. Pattern Recognit., Portland, OR,USA, Jun. 2013, pp. 708–715.
[10] M. Everingham, J. Sivic, and A. Zisserman, "Hello! My name is…Buffy—Automatic naming of characters in TV video," in Proc. 17thBrit. Mach. Vis. Conf., Edinburgh, U.K., Sep. 2006, pp. 899–908.
[11] J. Sang and C. Xu, "Robust face-name graph matching for movie character identification," IEEE Trans. Multimedia, vol. 14, no. 3, pp. 586–596,Jun. 2012.
[12] Y.-F. Zhang, C. Xu, H. Lu, and Y.-M. Huang, "Character identificationin feature-length films using global face-name matching," IEEE Trans.Multimedia, vol. 11, no. 7, pp. 1276–1288, Nov. 2009.
[13] M. Tapaswi, M. Bäuml, and R. Stiefelhagen, "'Knock! Knock! Who isit?' Probabilistic person identification in TV series," in Proc. 25th IEEEConf. Comput. Vis. Pattern Recognit., Providence, RI, USA, Jun. 2012,pp. 2658–2665.
[14] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal componentanalysis?" J. ACM, vol. 58, no. 3, pp. 1–37, 2011, Art. ID 11.
[15] Y. Deng, Q. Dai, R. Liu, Z. Zhang, and S. Hu, "Low-rank structurelearning via nonconvex heuristic recovery," IEEE Trans. Neural Netw.Learn. Syst., vol. 24, no. 3, pp. 383–396, Mar. 2013.
[16] K. Q. Weinberger and L. K. Saul, "Distance metric learning for largemargin nearest neighbor classification," J. Mach. Learn. Res., vol. 10,pp. 207–244, Feb. 2009.
[17] C. Shen, J. Kim, and L. Wang, "A scalable dual approach to semidefinite metric learning," in Proc. 24th IEEE Conf. Comput. Vis. PatternRecognit., Colorado Springs, CO, USA, Jun. 2011, pp. 2601–2608.
[18] B. McFee and G. Lanckriet, "Metric learning to rank," in Proc. 27thInt. Conf. Mach. Learn., Haifa, Israel, Jun. 2010, pp. 775–782.

[19] S. Andrews, I. Tsochantaridis, and T. Hofmann, "Support vectormachines for multiple-instance learning," in Proc. 16th Annu.Conf.Neural Inf. Process. Syst., Vancouver, BC, Canada, Dec. 2003,pp. 65–72.
[20] M.-L. Zhang and Z.-H. Zhou, "M3MIML: A maximum margin methodfor multi-instance multi-label learning," in Proc. 8th IEEE Int. Conf.Data Mining, Pisa, Italy, Dec. 2008, pp. 688–697.
[21] T. Cour, B. Sapp, C. Jordan, and B. Taskar, "Learning from ambiguouslylabeled images," in Proc. 22nd IEEE Conf. Comput. Vis. PatternRecognit., Miami, FL, USA, Jun. 2009, pp. 919–926.
[22] E. Elhamifar and R. Vidal, "Sparse subspace clustering," in Proc. 22nd IEEE Conf. Comput. Vis. Pattern Recognit., Miami, FL, USA, Jun. 2009,pp. 2790–2797.
[23] C. Lu, J. Feng, Z. Lin, and S. Yan, "Correlation adaptive subspacesegmentation by trace Lasso," in Proc. 12th IEEE Int. Conf. Comput.Vis., Sydney, VIC, Australia, Dec. 2013, pp. 1345–1352.
[24] S. Xiao, M. Tan, and D. Xu, "Weighted block-sparse low rank representation for face clustering in videos," in Proc. 13th Eur. Conf. Comput.Vis., Zürich, Switzerland, Sep. 2014, pp. 123–138.
[25] Z. Lin, R. Liu, and Z. Su, "Linearized alternating direction method withadaptive penalty for low-rank representation," in Proc. 24th Annu. Conf.Neural Inf. Process. Syst., Granada, Spain, Dec. 2011, pp. 612–620.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY