iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** VI **Month of publication:** June 2023

DOI: https://doi.org/10.22214/ijraset.2023.53751

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: 7.538
Volume 11 Issue VI Jun 2023- Available at www.ijraset.com

Music Recommendation Based on Face Emotion Recognition

Pawan Kumar[1], Deepanshi Mittal[2], Gudiya Sakshi[3], Anamay Seereesh Pandey[4]
[1]*Assistant Professor,* [2, 3, 4]*Student, Department of Information Technology, IMS Engineering College Ghaziabad India*

Abstract: ***We endorse a brand new method for gambling tune routinely the use of facial emotion Maximum of the present processes contain playing songs manually, the use of wearable computing devices, or classifying based totally on audio capabilities.Alternatively, we advocate alternating guide sorting and gambling. We have used a Convolutional Neuralcommunity for emotion detection. For song pointers,Pygame & Tkinter are used. Our proposed machine tends to reduce the computational time worried in acquiring the results and the overall cost of the designed gadget, thereby increasing the device's standard accuracy. trying out of the gadget is carriedout at the FER2013 dataset. Facial expressions are captured with the use of an in-built digicam. Characteristic extraction is carried out on entering face images to hit upon feelings together with satisfied, indignant, sad, marvel, and neutral. A routine song playlist is generated by figuring out thecurrent emotion of the user. It yields higherperformance in phrases of computational time, as compared to the set of rules within the current literature.***
Keywords: Face Recognition, Convolutional Neural Network, Tkinter, Music, Emotion detection, Player, Camera, Feature extraction, Pygame.

I. INTRODUCTION

Various studies in latest years admit that people reply andreact to music and this track has a high influence on the interest of the human mind. In a single examination of the reasons why humans pay attention to music, researchers discovered that music performed a crucial role in touching on arousal and temper. Two of the most important functions of the song are its far potential members rated to help them acquire a great temper and turn out to be more self-aware. Musical options were validated to be fantastically associated with persona trends and moods [1].

The meter, timbre, rhythm, and pitch of tune are controlled in regions of the brain that influences emotions and temper [2]. The interplay between individuals may bea primary aspect of lifestyle. It is famous for perfect details and much of facts among humans, whether or notthey're in the shape of frame language, speech, facial expression, or feelings [3]. Nowadays, emotion detection is considered the maximum crucial approach used in lots of packages together with clever card programs, surveillance, photograph database research, criminal, video indexing, civilian applications, protection, and adaptive human-computer interface with multimedia environments.

With the growth in an era of digital signal processing andother effective function extraction algorithms, automatic emotion detection in multimedia attributes like tunes or movies is growing hastily and this gadget can play an important function in lots of capability programs like human-computer interaction systems and music enjoyment. We use facial expressions to advise a recommender device for emotion popularity which can detect a person's emotions and advocate a list of suitable songs [13-24]. The proposed system detects the feelings of someone, if the character has a terrible emotion, then acertain playlist will be shown that consists of the maximum associated types of tune so one can decorate his mood. And if the emotion is nice, a selected playlist could be offered which incorporates unique styles of music so one can inflate the positive feelings [4].

The dataset we are going to used for emotion detection is from Kaggle's facial features reputation [5]. The dataset for the track participant has been constituted of Bollywood Hindi songs. The implementation of facial emotion detection is done with the usage of Convolutional. Neural network which offers about95.14% of accuracy [2].

II. LITERATURE REVIEW

The evaluation is executed to get insights into the methods, and their shortcoming which we can conquer. Aliterature evaluation is a textual content of a scholarly paper, which includes the modern understanding at the side of high-quality findings, in addition to theoretical andmethodological contributions to a selected subject matter.The latent features of people which can provide inputs toany gadget in numerous approaches have introduced the attention of numerous novices, scientists, engineers, etc. from all over the globe.

©IJRASET: All Rights are Reserved

The current intellectual nation of the person is supplied through facial expressions. Most of the time we use nonverbal clues like hand gestures, facial expressions, and tone of voice to express emotions in interpersonal conversation. Preema et al [6] stated that it is very time- eating and tough to create and manipulate a massiveplaylist. The paper states that the `track participant itself selects a track in line with the current temper of the user. The utility scans and classifies the audio files in keeping with audio functions to produce temper-based playlists. The utility makes use of the Viola-Jonas set of rules whichis used for face detection and facial feature extraction. Support Vector machine (SVM) changed into used inside the category extracted capabilities into five fundamental frequent emotions like anger, joy, surprise, unhappiness, and disgust.

Yusuf Yaslan et al. proposed an emotion-primarily basedsong recommendation system that learns the consumer's emotion from signals obtained via wearable computing devices that are incorporated with galvanic skin reaction (GSR) and photoplethysmography (PPG) physiological sensors of their paper [3]. Feelings are a basic part of human nature. They play a crucial position for the duration of life. In this paper, the emotion popularity trouble is taken into consideration as arousal and valence prediction from multi-channel physiological alerts. In [7]Ayush Guidel et al stated that humans' kingdom of thoughts and current emotional temper may be easily located thru their facial expressions. This system was developed by using taking primary feelings (happiness, sadness, anger, exhilaration, marvel, disgust, worry, and neutrality) into attention. Face detection in this project became carried out with the aid of using a convolutional neural community. The tune is normally instructed as a "language of feelings" throughout the planet.

The paper proposed via Ramya Ramanathan et al [1] conveyed the smart song player the usage of emotion popularity. Emotions are a totally basic part of human nature. They play the maximum vital position at some stage in existence. Human feelings are intended for sharing feelings and mutual information. The person's neighbourhood music selection is to begin grouped basedtotally on the emotion conveyed via the album. That is often calculated taking into account the tune's lyrics. Thepaper especially makes a strong point about the methodologies available for detecting human emotions for growing emotion-based total track players, the approach a song participant follows to stumble on humanemotions, and the way it is good to use the proposed machine for emotion detection. It additionally offers a brief idea about our systems running, playlist generation, and emotion classification.CH Radhika et al [8] counselled guide segregation of a playlist and annotationof songs, following the current emotional nation of a user,as an exertions-in depth and time-eating challenge. Numerous algorithms were proposed to automate thisway. But, the prevailing algorithms are slow, increase theoverall cost of the gadget with the aid of using extra hardware (e.g., EEG systems and sensors), and function much much less accurately. The paper gives an algorithmthat mechanically does the process of generating a playlistof audio, based totally on the facial expressions of someone, for rendering salvage of time as well as hard work, invested in acting this method manually. The algorithm given within the paper directs toward reducing the overall computational time and the fee of the designed gadget. It moreover pursues developing the accuracy of the machine layout. The machine's facial expression popularity module is validated by comparing it to a dataset that is each user-based and consumer-unbiased.

III. BACKGROUND WORK

It is often puzzling for a person to decide which music he/she needs to concentrate on from a big collection of existing options. There were numerousproposal frameworks to be had for troubles likemusic, dining, and shopping depending upon thetemper of the person. The principle objective of our song advice device is to provide guidelines to the customers that fit their person options. The analysis of the facial expression/person's emotion might alsolead to expertise in the contemporary emotional or intellectual country of the consumer. Music and videos are one region where there's a massive risk toprescribe plentiful alternatives to clients in light of their tendencies and additionally recorded information. It is widely recognized that human beings employ facial expressions to express extra absolutely what they want to say and the context in which they meant their words. Greater than 60 percent of the customers believe that at a sure point in time the range of songs found in their songs libraryis so big that they're unable to determine the music which they should play. Via developing an advice machine, it may assist a person to decide which songone ought to pay attention to assisting the person to reduce his/her strain ranges.

The consumer would no longer waste any time looking or to appearance up for songs and the exceptional track matching the consumer's mood is detected, and songs would be shown to the consumer in keeping with his/her mood. The photo of the person is captured with the help of a webcam. The person's image is taken afterwhich as per the temper/emotion of the consumer appropriate music from the playlist of the user is proven to match the person's requirement. The person's photograph is taken after which as in step with the temper/emotion of the consumer the correcttrack from the playlist of the consumer is played matching the person's requirement.

©IJRASET: All Rights are Reserved

The gadget has efficiently been capable of seizing the emotion of a consumer. It has been examined in a real-time environment for this predicate. It has to be, however, examined in special lighting situationsto determine the robustness of the evolved gadget the system has also been able to seize brand new photographs of the consumer and it should replace itsclassifier and training dataset. The machine was designed using the facial landmarks scheme and is examined below diverse eventualities for the result that might be received. It's far visible that the classifier has an accuracy of more than 80 percent formost of the take-a-look-at cases, which is quite properly accurate in phrases of emotion type. It can additionally be seen that the classifier can correctly expect the expression of the user in a real-time scenario while examining the stay of a person.

IV. PROPOSED APPROACH

The proposed machine advantages us to give interplay between the person and the song player. The motive of the device is to seize the face well with the digital camera. Captured pics are fed into the Convolutional Neural network which predicts the emotion. Then Emotion derived from the captured picture is used to geta playlist of songs. The primary intention of our proposeddevice is to provide a track playlist mechanically to exchange the user's moods, which can be satisfied, sad, natural, or amazed. The proposed device detects emotions, if the topic features a terrible emotion, then aspecific playlist is going to be presented thatincorporates the most suitable types of music with the purpose to enhance the temper of the individualundoubtedly. Song advice primarily based on facial emotion reputation carries four modules.

1) Real-Time seize: on this module, the gadget is tocapture the face of the consumer efficaciously
2) Face reputation: right here it'll take the user's face as enter. The convolutional neural community is programmed to evaluate the capabilities of the consumer image.
3) Emotion Detection: in this segment extraction ofthe functions of the person image is executed tostumble on the emotion and depending on the consumer's feelings, the machine will generate captions.
4) Song recommendation: the tune is usually recommended using the advice module to the user using mapping their feelings to the temper sort of the track.

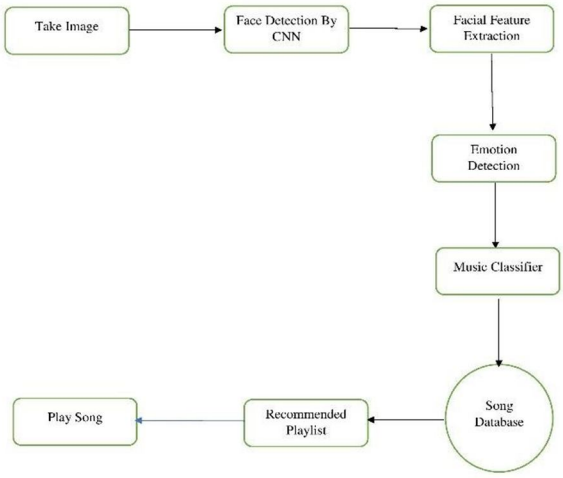


Figure 1. Architectural diagram of proposed

V. METHODOLOGY

We built the Convolutional Neural community version using the Kaggle dataset. The database is FER2013 which is split into elements training and checking out the dataset. The training dataset consists of 24176 and the testing dataset incorporates 6043 pictures. There are 48x48 pixel grayscale pix of faces inside the dataset. Every image in FER-2013 is categorized as one of 5 feelings: happy, unhappy, indignant, marvel, and impartial.

The faces are robotically registered in order thatthey're more or much less focused in every photograph and take up approximately the same amount of space. Theimages in FER-2013 comprise each posed and unposed headshot, which can be in grayscale and 48x48 pixels.

The FER-2013 dataset was created by way of accumulating the results of a Google image seek of each emotion and synonyms of the feelings. FER systems being educated on an imbalanced dataset may perform properly on dominant feelings consisting of happy, unhappy, irritated, neutral, and amazed however they perform poorly at the beneath-represented ones like disgust and worry. Typically, the weighted-softmax loss approach is used to handle this trouble by using weighting the loss term for every emotion magnificence supported using its relative share in the education setting. But, this weighted-loss approach is predicated on the softmax lossfeature, that's stated to effortlessly force functions ofvarious classes to stay apart without being attentive to intra-class compactness. One powerful strategy to address the matter of softmax loss is to use an auxiliary loss to teach the neural community. To treat lacking and Outliers values we've got used a loss feature named express cross-entropy. For each iteration, a specific loss feature is hiredto gauge the mistake cost. So, to treat lacking and Outliersvalues, we've used a loss characteristic named express cross-entropy.

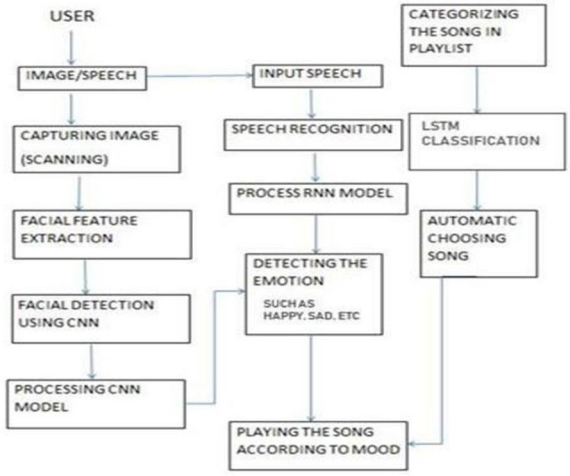


Figure.1. Flowchart Of Methodology

A. Emotion Detection Module

1) Face Detection

Face detection is one of the packages that's considered beneath laptop imaginative and prescient technology. That is the system wherein algorithms areevolved and skilled to correctly locate faces or gadgets in object detection or associated gadgets in pictures. This detection may be real-time from a video frame or pics. Face detection uses such classifiers, which can be algorithms that stumble on what is either a face (1) or now not a face (zero) in a photo. Classifiers are educated to detect faces in the usage of numbers of pix to get greater accuracy. Opencv makes use of forms of classifiers, LBP (neighbourhood Binary pattern), and Haar Cascades.A Haar classifier is used for face detection where theclassifier is trained with pre-defined various face facts which allows it to come across one-of-a-kind faces appropriately. The primary goal of face detection is to spot the face in the frame with the aidof reducing external noises and other factors. It is a device gaining knowledge of-based the total method in which the cascade characteristic is skilled with a group of entered documents. It is supported by the Haar Wavelet approach to investigate pixels in the image into squares through characteristics . This usesmachine getting-to-know strategies to induce an excessive diploma of accuracy from what is known as "education facts"



Figure.2 Face Detection

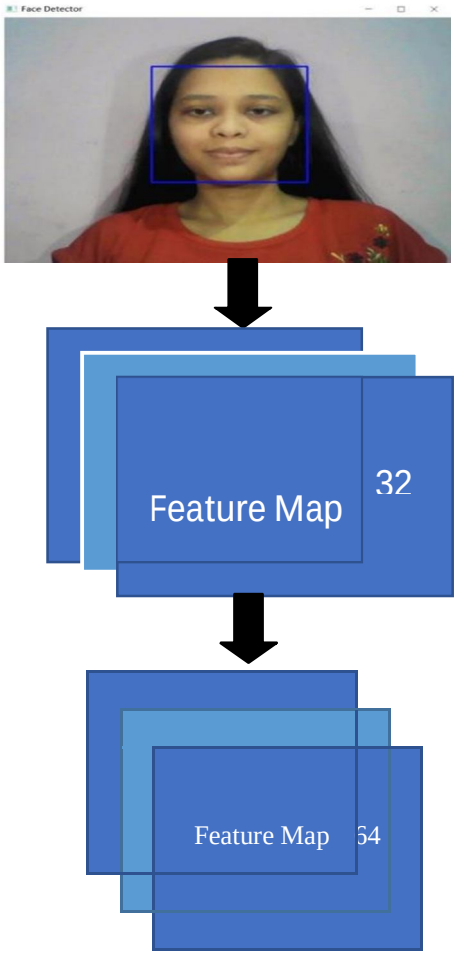
©IJRASET: All Rights are Reserved

2) Feature Extraction

At the same time as performing characteristic extraction, we treat the pre-educated network which is a sequential model as an arbitrary characteristic extractor. Permitting the entered picture to pass on itahead, stopping at the pre-distinct layer, and taking the outputs of that layer as our features. Starting layers of a convolutional network extract excessive degree functions from the taken image, so use only afew filters. As we make in addition deeper layers, weincrease the quantity of the filters to twice or three times the size of the filter out of the preceding layer. Filters of the deeper layers advantage greater capabilities but are computationally very in-depth. In doing this we utilized the robust, discriminative features learned via the Convolution neural network.The outputs of the version are going to be characteristic maps, which can be an intermediate representation for all layers after the very first layer. Load the entered photograph for which we want to view the characteristic map to recognize which capabilities had been prominent to classify the picture. Characteristic maps are obtained by using applying Filters or function detectors to the entered image or the characteristic map output of the previous layers. Feature map visualization will offer perception into the indoor representations for specific. In position for each of the Convolutional layers inside the model.

3) Emotion Detection

Convolution neural network architecture applies filters or feature detectors to the entered picture to get the feature maps or activation maps using the Relu activation function. Feature detectors or filters assist in figuring out various functions present within the photograph which include edges, vertical strains, horizontal lines, bends, etc. After that pooling is implemented over the characteristic maps for invariance to translation. Pooling is expected on the idea that once we trade the input with the aid of a hint quantity, the pooled outputs don't alternate. We can use any of the pooling from min, average, or max. However, max-pooling affords higher performance than min or common pooling. Flatten all the entries and give those flattened inputs to a deep neural network which can be output to the elegance of the item.



ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue VI Jun 2023- Available at www.ijraset.com

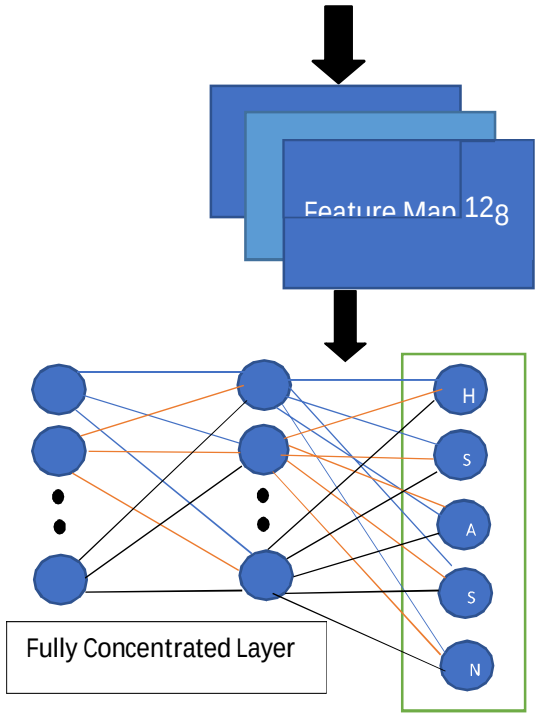


Figure.3 Neural Network Architecture

B. Music Recommendation

1) Song Database: We created a database for Bollywood Hindi songs.It includes one hundred to a hundred and fifty songs steady with emotion. As anybody apprehends music is undoubtedly concerned with improving our mood. So, think a person is unhappy then the gadget will propose this type ofmusic playlist which motivates her or him and byway of this automatic mood can be pleased.

2) Song Recommendation: By the use of the emotion module real-time emotion of the user is detected. This could supply the labels like happy, sad, angry, marvel, andneutral. The use of the os.listdir() method in python we linked those labels with the folders of the songsdatabase which we have created. Desk 1 indicates the list of songs. This method of os.listdir() is usedto get the listing of any record inside the specific directories.

Desk.1. Songs Database

EMOTION	SONGS
Happy	Track.1 “Badtameez Dil”
	Track.2 “Happy Hours”
	Track.3 “Matargashti”
Sad	Track.1 “Tujhe Bhula Diya”
	Track.2 “Do Pal”
	Track.3 “Main Jahan Rahoon”
Angry	Track.1 “Dushman Na Kare”
	Track.2 “Thukra Ke”
	Track.3 “Junoon”
Marvel	Track.1 “Mitwa”
	Track.2 “Theher Ja”
	Track.3 “ Kun Faya Kun”
Neutral	Track.1 “Paisa”
	Track.2 “Hello”
	Track.3 “Jee Le”

©IJRASET: All Rights are Reserved

This can bring about the advocated playlist for theconsumer in the GUI of the tune participant by displaying captions in line with detected feelings.We've got used a library known as Pygame for gambling the audio as this library helps play diverse multi-media formats like audio, video, andso forth. Functions of this library includingplaysong, pausing, resume song, and stop song areused to work with the song participant. Variables like playlist, song status, and root are used for storing the call of all songs, storing the reputationof presently lively songs, and for the principal GUIwindow respectively. For developing the GUI,we've used Tkinter.

VI. RESULT AND DISCUSSION

We evaluated some of the studies which use help vector gadgets (SVM), severe learning machines (ELM), and convolutional neural communities [12]. Desk 2 suggests the comparison of related algorithms. Corresponding algorithms andaccuracy values are given for every look. The use of a Convolutional Neural network improves the performance of emotion detection accuracy.

Algorithm	SVM	ELM	CNN
Validation Accuracy	0.66	0.62	0.95
Testing Accuracy	0.66	0.63	0.71

Desk.2. Accuracies Of Algorithm

VII. CONCLUSION

Thorough evaluation present day the literaturetells that there are numerous methods to put into effect song Recommender machines. A have a look at modern-day strategies proposed with theAid of preceding scientists and developers changed into completed. Primarily based on the findings, the objectives latest our gadget had been fixed. Asthe electricity and blessings of modern-day AI-powered applications are trending, our challenge might be a trending generation usage. On this machine, we provide an outline of ways tune can have an effect on the consumer's temper and how to pick out the proper music tracks to enhance theperson's moods. The carried-out system can detectthe person's feelings. The emotions that the machine can come across have been glad, unhappy, anger, impartial, or amazed. Afterdetermining the user’s emotion, the proposed system provided the person with a playlist that carries music matches that detected the temper. Processing a huge dataset is memory in addition toCPU in depth. This will make improvement more tough and appealing. The reason is to create this application in the most inexpensive possible way and also to create it below a standardized device. Our track advice gadget based on facial emotion reputation will lessen the efforts of modern-day customers in growing and coping with playlists.

VIII. FUTURE WORK

This gadget, although functioning, does have scope for improvement within the destiny. There are various components of the software that can bechanged to provide better results and a smoother universal enjoyment for the person. A number of these that an alternative method, based on additional emotions which can be excluded in ourmachine as disgust and fear. This emotionprotected assisting the gambling of tracks automatically. The future scope of the device could style a mechanism that might be beneficial in music therapy remedies and assist the tunetherapist to deal with patients suffering from mental stress, anxiety, acute melancholy, and Trauma. The modern device no longer carries nicely in extremely horrific light conditions and the negative digicam decision thereby gives the possibility to add a few capabilities as a solution inside the destiny.

REFRENCES

[1] Ramya Ramanathan, Radha Kumaran, Ram Rohan R, Rajat Gupta, and Vishalakshi Prabhu, anintelligent music player based on emotion recognition, 2nd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2017. https://doi.org/10.1109/CSITSS.2017.8447743

[2] Shlok Gilda, Husain Zafar, Chintan Soni, Kshitija Waghurdekar, Smart music player integrating facial emotion recognition and music mood recommendation, Department of ComputerEngineering, Pune Institute of Computer Technology, Pune, India, (IEEE),2017. https://doi.org/10.1109/WiSPNET.2017.8299738

[3] Deger Ayata, Yusuf Yaslan, and Mustafa E.Kamasak, Emotion-based music recommendationsystem sensors, IEEE transactions on consumerelectronics, vol. 14, no. 8, May 2018 https://doi.org/10.1109/TCE.2018.2844736

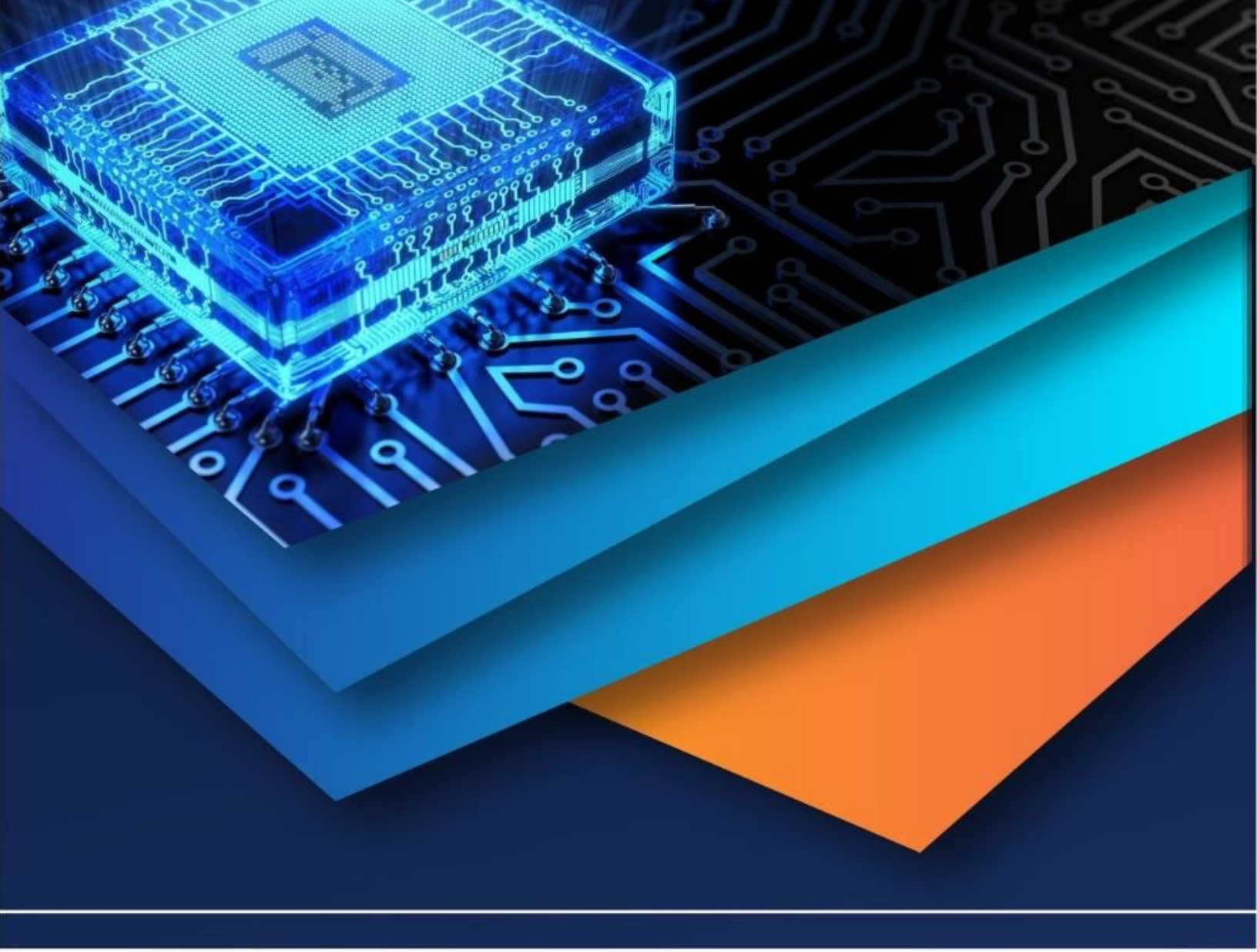
[4] Ahlam Alrihail, Alaa Alsaedi, Kholood Albalawi, Liyakathunisa Syed, Music recommender system for users based on emotion detection through facial features, Department of Computer Science Taibah University, (DeSE), 2019. https://doi.org/10.1109/DeSE.2019.00188

[5] Research Prediction Competition, Challengesin representation learning: facial expression recognition challenges, Learn facial expression from an image, (KAGGLE)

[6] Preema J.S, Rajashree, Sahana M, Savitri H, Review on facial expression-based music player, International Journal of Engineering Re search & Technology (IJERT), ISSN-2278-0181, Volume 6, Issue 15, 2018.

[7] AYUSH Guidel, Birat Sapkota, KrishnaSapkota, Music recommendation by facialanalysis, February 17, 2020.
[8] CH. sadhvika, Gutta.Abigna, P. Srinivas reddy, Emotion-based music recommendationsystem, Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad;International Journal of Emerging Technologies and Innovative Research (JETIR) Volume 7, Is- sue 4, April 2020.
[9] Vincent Tabora, Face detection using OpenCVwith Haar Cascade Classifiers,Becominghuman.ai,2019.
[10] Zhuwei Qin, Fuxun Yu, Chenchen Liu, XiangChen. How convolutional neural networks see theworld - A survey of convolutional neural networkvisualization methods. Mathematical Foundationsof Computing, May 2018.
[11] Ahmed Hamdy AlDeeb, Emotion-Based Music Player Emotion Detection from LiveCamera, ResearchGate, June 2019.
[12] Frans Norden and Filip von Reis Marlevi, A Comparative Analysis of Machine Learning Algorithms in Binary Facial ExpressionRecognition, TRITA-EECS-EX-2019:143.
[13] P. Singhal, P. Singh and A. Vidyarthi (2020)Interpretation and localization of Thorax diseases using DCNN in Chest X-Ray. Journal of Informatics Electrical and Elecrtonics Engineering,1(1), 1, 1-7
[14] M. Vinny, P. Singh (2020) Review on the Artificial Brain Technology: BlueBrain. Journal of Informatics Electrical and Electronics Engineering,1(1), 3, 1-11.
[15] A. Singh and P. Singh (2021) Object Detection. Journal of Management and Service Science, 1(2), 3, pp. 1-20.
[16] A. Singh, P. Singh (2020) Image Classification: A Survey. Journal of Informatics Electrical and Electronics Engineering,1(2), 2, 1- 9.
[17] A. Singh and P. Singh (2021) License Plate Recognition. Journal of Management and ServiceScience, 1(2), 1, pp. 1-14

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)