Face Expression and Emotion Detection by using Machine learning and Music Recommendation

Prince Kumar[1],Gouri Sankar Mishra[2],Tarun Maini[3],Pradeep Kumar Mishra[4], Shubham Dubey[5], Shivam Sharma[6]

[1,4,5,6]Department of Computer Science & Application, Sharda School of Engineering & Technology, Sharda University, Greater Noida, UP, India
[2,3]Department of Computer Science & Engineering, Sharda School of Engineering & Technology, Sharda University, Greater Noida, UP, India.

1 prince88735@gmail.com ,2gourisankar.mishara@sharda.ac.in , 3 tarun.maini@sharda.ac.in, 4 pradeepkumar.mishra@sharda.ac.in, 5 subhamdubey6395@gmail.com ,6 samsharma2349@gmail.com

ABSTRACT

Most of us listen to music to feel emotions. Your negative mood might be lifted by music. Currently existing music systems let you listen to chosen music and suggest songs in categories depending on your interests or the tastes of other users. Music fans cannot completely depend on such methods and therefore do not prefer to listen to music on the station or online while such sound systems are not created with the emotions elicited in mind. In this work, we provide a music system based on sentiment. Our Raspberry Pi-based system plays tunes based on the ambiance of the room using a speaker, a microphone, and a Raspberry Pi. The emotion of the recorded background sounds is assessed using a classification issue based on machine learning. For this categorization, we make use of a simple Bayesian classifier. Using the song's Bits per Minute pace to identify songs with comparable emotional content.

Keywords: Face-extraction, music suggestion, emotion recognition, and real-time image capture.
DOI Number: 10.48047/NQ.2022.20.20.NQ109290 **NeuroQuantology2022;20(20):2945-2954**

1. Introduction

By identifying and recording the user's emotions in real time, this notion suggests music to the user.We presented a method to categorise various types of music into distinct moods, such as joyful, sad, furious, etc. Previous techniques were employing collaborative techniques that used user data from prior sessions to select music. However, these techniques take a lot of human work. Emotion-Based-music-player It is a music player that uses Chrome as its front-end and a machine learning algorithm written in Python to recognise emotions on the user's face. The user will be shown/recommended songs based on the mood that has been identified for them.

In this application, a person's picture is taken utilizing a real-time device that can access the nearby equipment. Other than that, we have some common features like a queue playlist so that we can have a personal playlist, and the last one is random. It uses the python Eel library so that it can pick a random song without any order. Also, based on the image that was shot, the database data sets previously saved on the local device are compared, and after processing, it determines the user's current mood in numerical form

dependent on how the music is going to be played. For this, we used tools like OpenCV, EEL, numpy, etc.
This method focuses mostly on suggested Music and has evolved an essential tool for reducing stress in modern society. Since facial expressions frequently convey emotion, we use faces as our major source of information for identifying emotion. Next, We provide music that can alter a user's mindset in line with that user's mood.

2. Literature review

2.1. System for Detecting Faces and Recognizing Expressions on the Face

Anagha S. Dhavalikar[1] suggested a technique for automatically recognising facial expressions. This system consists of three phases. Facial recognition, Expression recognition, followed by Feature extraction.The InitialRGB Color Model is used for face detection, adjusting the lighting while acquiring a face and morphological surgeries to keep the desired face, such as the face's lips and eyes.This System is also utilisedActive Appearance Model Technique, or AAM, is used to extract face features. In this approach,The model facial angularities, including the lips, br0ws, and eyes, are located, with a file holding data details about the model points that were identified is produced. The method also detects faces and uses input of an expression to determine how the AAM Model should change.

2.2. Bezier Curve Fitting for Emotional Identification from Face Expression Analysis

Bezier curve fitting was the basis for the approach provided by Youngseop Kim, Woori Han, and Yong-Hwan Lee[2]. The first stage in this system's method for determining facial expression and In order to validate the facial expression of certain characteristics in the region of interest, the second step is to identify and analyze the facial landmarks from the original photo input. determining the position of the lips and eyes on the face as well as the angle of the face, feature maps were employed after the initial step of face identification, which used color still images considering skin color pixels and initiated spatial filtering. In the process of applying a Bezier curve to the eye and mouth, this approach first extracts the targeted area before the feature map's points for extraction. This method employs training and measurement of the Hausdorff separation between the input face picture and the data image using a Bezier curve in order to comprehend emotion.

2.3. Using animated mood images to suggest music

A technique for recommending music employing animated mood images was proposed by ArtooLeptinemia and JukkaHolm[3]. Using a library of photos, the user of this system may obtain music recommendations based on the genre of each image**.** The Nokia Research Center created this technique for making music recommendations. Audio signal processing and textual meta tags are used in this system to describe the genre.

2.4. Utilizing emotion identification from facial expressions in human-computer interaction.

A.Pruski et al., C. Maaoui et al., and F. Abdat[4]. They suggested a completely automated facials emotion and identification system based on three steps: face recognition, facial feature categorization of facial expressions and extraction. This methodology coupled the Shi and Tomasi method with an anthropometric model to identify the facial feature points. This methodology uses a version of 21 distances to characterize facial features from a neutral face and classifies data using SVM.

2.5. Music Suggestions Depend on Emotion By association, I discovered film music.

Suh-Yin Lee et al. and Fang-FeiKuo et al .[5] The growth of music suggestions for customers is a result of the spread of digital music. The consumers' preferences for music are the basis for the current recommendation methods. Nonetheless, there are occasions when selecting music based on the mood is necessary. Using association learning from cinema music, We provide a unique approach for recommending music based on emotions. In this work. In order to uncover associations between emotions and musical qualities, we examined musical feature extraction and modified the affinity graph. According to experimental findings, the suggested technique averages 85% accuracy.

2.6. Interactive Music Search and Recommendation Based on Mood

According to John O'Donovan et al. and Ivana Andjelkovic et al. [6],On increasing prediction and ranking, recommender system research has been

heavily concentrated. The importance of other aspects of the suggestions, such as accessibility, flexibility, and overall user experience, has, however, been highlighted by recent studies. On the basis of these features, we propose MoodPlay, a hybrid music recommendation system with a user-friendly interface that combines content- and mood-based filtering. We walk users through using MoodPlay to search music files by secret emotional aspects and how to blend users input with forecast made from a prior user profile when making recommendations. Findings from a user research (N=240) that looked at 4 situations with various visibility, engagement, and control levels are discussed.

2.7. An Reliable Face Expression-Based Music Playlist Generation Algorithm

AnukritiDureha and colleagues [7]. He recommended labor- and time-intensive manual playlists segmentation and music annotation depending on the user's present state of mind.To automate this procedure, many algorithms have been suggested. The present algorithms, however, are less precise, require more equipment (such as EEG sensors and equipment), which drives up the cost of the system as a whole, and are wasteful. the method of creating an audio album from a participant's facial gestures to save time and labor-intensive manual labor. The algorithm put forward in this research aims to cut down on both the system's total cost and calculation time. It also seeks to improve the proposed system's accuracy. The proposed algorithm's facial expression recognition module is tested against user-dependent and user-independent datasets to ensure its accuracy.

2.8. Improving Character Details and Emotional States in Music Advisory Systems

Markus Schedl et al. and Bruce Ferwerda et al. [8]. made the suggestion that the basic study hypotheses be improved by including personality and emotional states into music choices.We think that by taking these psychological elements into account, the recommendation's accuracy may be improved. The system focuses on the relationship between a person's personality and how they utilize music to control their emotional states[9].

It's essential to be able to discern an individual's emotions from their face.In order to capture the essential data from the a face of a person.Among other things, input input may be used to extract information that may be used to estimate a person's mood.Songs are created using the "feeling" that is acquired from the preceding input. A playlist that is appropriate for a certain person's emotional qualities may be made with less effort spent manually sorting songs into numerous categories. In order to produce a playlist that meets the specified criteria,Upon scanning and understanding the data, the Facial Expression Based Music Player. Our proposed method focuses on recognising constructing an emotion-based music player using personal emotions. It explains how our music player detects human emotions, how other music players currently on the market sense emotions, and how to use our technology for emotion detection to its fullest potential. Also, a brief explanation of playlist generation, emotion categorization, and the operation of our algorithms is provided.We utilized the pycharm tool for analysis in this project[10].

The results of research are divided into two phases:

1. Using Python, create a programme that can identify a user's emotion from their expression.
2. The music will be played based on the user's preferences if Python-code is added to the web-service. statement.

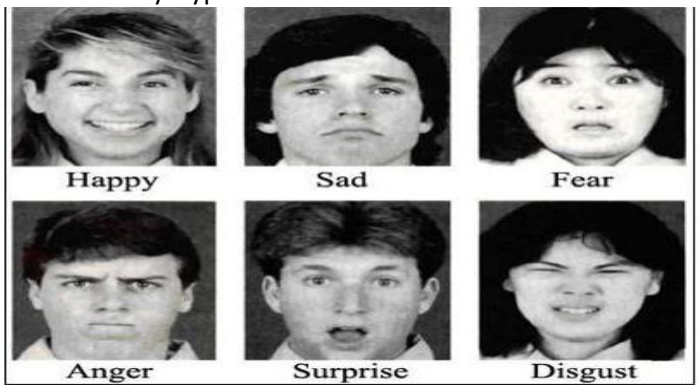

Fig-1. Various of face Expression

3. Proposed methodology

In web camera captures the Particulars face. Frames from the captured video are created. Utilizing preprocessing, the webcam picture is used to transform the facial expression into a series of Actions Unites (AUs).The Facial Action Coding System makes use of combinations of the 64 AUs characterizes every face emotion. Following feature extraction, the faces' emotions—such as happiness, anger, sadness, and surprise—are categorized. They are connected with the web services. They might be SAAS, IAAS, or PAAS. The music is played based on the emotions that are recognised, and the feelings are communicated[11].

3.1. Fisher Face Algorithm

This image processing system employs the principle component analysis (PCA) approach to reduce the size of the face space before obtaining the feature of the image characteristics using the fisher's linear discriminant (FDL) or local discriminant analysis (LDA) methods. Because it optimizes the distinction between classes throughout the training phase, we specifically adopt this approach. While the minimal Euclidean technique is used for matching faces, this approach aids in picture identification and helps us categorize facial expressions that suggest user mood[20].

3.2. Haar Cascade Algorithm

The Haar Cascade Algorithm is a machine learning tool for classifying the various elements of a taken picture. The major use is object detection. Several steps of collection from weak learners are used in the cascade classifier. These weak classifiers, also known as boosting classifiers, are the most basic kind of classifiers. If the label has a positive range, the process advances to the step where the outcome is shown. They recognise the photos in accordance with the labels, which has both good and bad aspects. On various levels, they have a group of positive photos placed over the unfavorable ones. Images with more clarity and moreof them are favored since they produce better outcomes.

In this case, the object in the image is found using the Haarcascade frontal face default.xml algorithm. Nose, eyes, ears, and lips are the objects in this face[13]. The frontal face is detected using the open cv-designed Haarcascade. Also, it has the ability to recognise the source's characteristics. It functions by superimposing negative pictures on top of positive ones that have been trained over them. Just the photos that we want our classifier to classify are contained in positive images. Negative images are made up of all other images, none of which contain the item we're looking for[14].

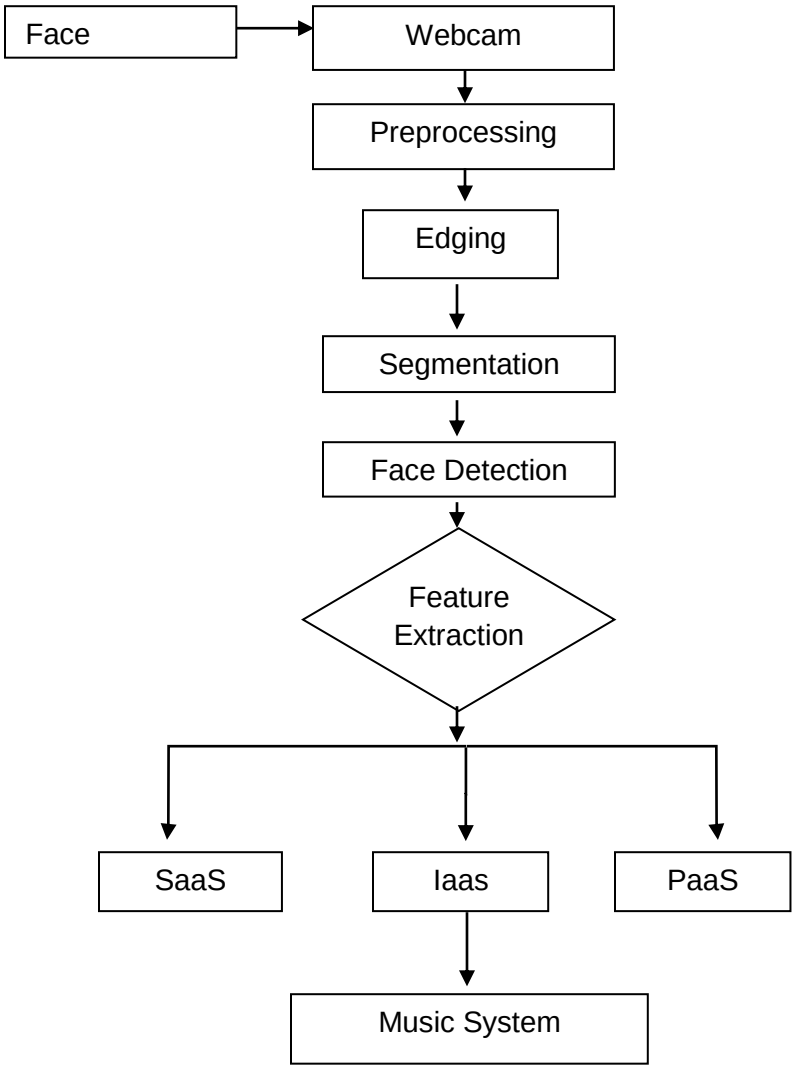

Fig-2 Schematic of the Emotion-Based Music Recommendation System's overall architecture.

4. Experiments and results

4.1. Face recognition

By eliminating extraneous sounds and other elements, the major goal of the face identification approaches to recognise the face in the picture(faces). The FACE DETECTION METHOD has the following steps:

1. Picture pyramid
2. Oriented gradients histogram
3. Linear Classifier

Using an image pyramid with several scales, the resulting data are divided into the sample image. Just extracting features while lowering noise and other parameters is how this approach is used.The Gaussian pyramid, also known as the lowpass filter image pyramid technique, subsamples the frame by decreasing its resolution and smoothing it. To get the intended result, a frame that resembles greater

resolution and a higher amount of smoothing than the original, the operation must be performed numerous times.

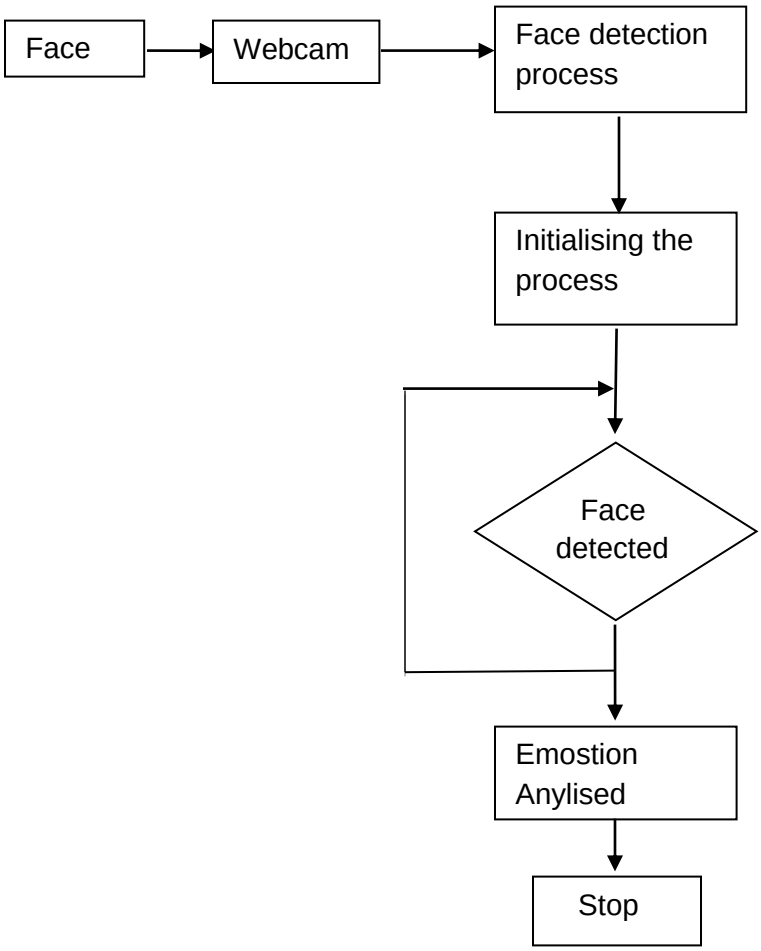

Fig-3. Flowchart for the Face Detection Module

The term "HOG" refers to a feature descriptor used when it comes to image processing that counts the instances of gradient orientations in a certain region of an image and is widely applied to the identification of objects in photographs.The basic goal of this approach is to use a collection of distributions of intensity gradients to characterize the face in the image.

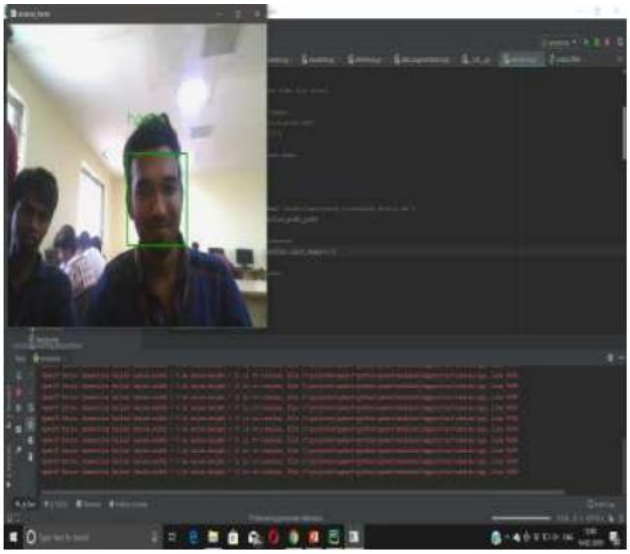

Fig-4. Facial recognition method

The final stage of the face detection procedure is linear classification. We only substituted using a linear predictor for SVM in order to shorten the computing set-up required for classification and, thus, provide a quicker face identification operation[15][16].

4.2. Classification of emotions

The image will have a bounding box overlayed over it once the faces has been properly recognised in order to extractable the ROI (face) for additional analysis. The 68 facial feature points will then be retrieved using the "Predictor" function, sometimes referred to as a script from the extracted ROI and save them in an array. A PCA reduction process will then be used to compress the data from the features array. Then remove leaving only the major components after removing any associated coordinates of the important sites. The 68x2 array of the data contains coordinates on the x- and y-axes for each of the 68 points. The array will be changed into a 1 column by 136 row vector. A series of photos and landmark maps for each image are used to train the face landmark extraction algorithm "Predictor[17]."

The methodology uses regression approaches taught with a gradient boosting approach to learn how to obtain the facial landmark map from a given face shot using just the pixel's intensity values indexing of each point. After the PCA reduction procedure, the data will be categorised [18]. Using a multiclass SVM with a linear kernel, the given data is compared with the stored data to decide which class (feeling) it belongs to. If any of the three emotions—anger, fear, or surprise—are noticed, a speed-reducing command will be sent to reduce the wheelchair's speed in order to keep the user safe[19].

4.3. Music suggestion

The webcam is used to capture the video before the frame is finished since the input is Captured in real-time. The processed framed photographs are classified using a hidden markov chain. For the goal of emotion categorization, all frame and pixel formats from the collected frames are taken into account. Each facial landmark's value is determined, and recorded for further use. Most classifiers are successful to a degree of 90–95%. such that even if the face changes as a result of outside influences.

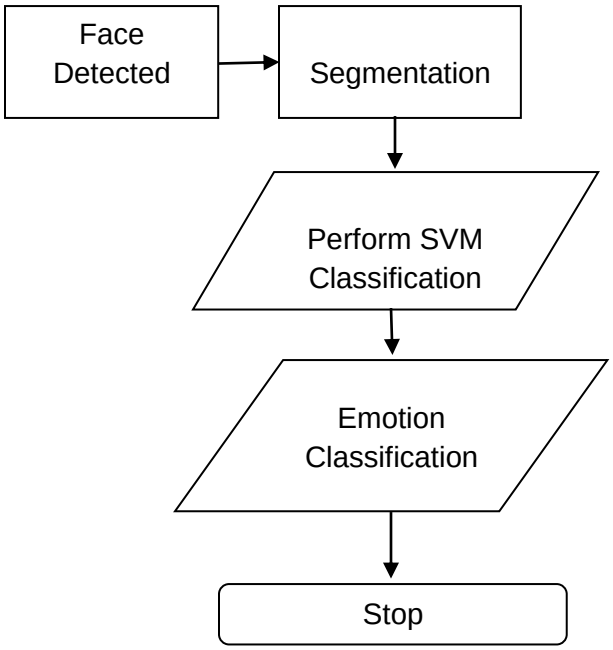

Figure No. 5: Emotion Classification Module Flow Diagram

The systems still has recognise the facial and the sentiment conveyed. The values retrieved and set are then utilized to determine the feelings,and the value of the received pixel is contrasted with the values contained as the code's threshold. From the user to the online service, values are transmitted. Based on the feeling experienced, the music is performed. Every song has a certain set of emotions.The appropriate music will play when the desired feeling is conveyed. Happy, angry, sad, and surprised are the four emotions that can be employed[15]. The music are played in line with the feelings that are detected; in other words, when the joyous sensation is recognised, the songs allocated for that particular emotion are played, which is also true of the other emotions[20].

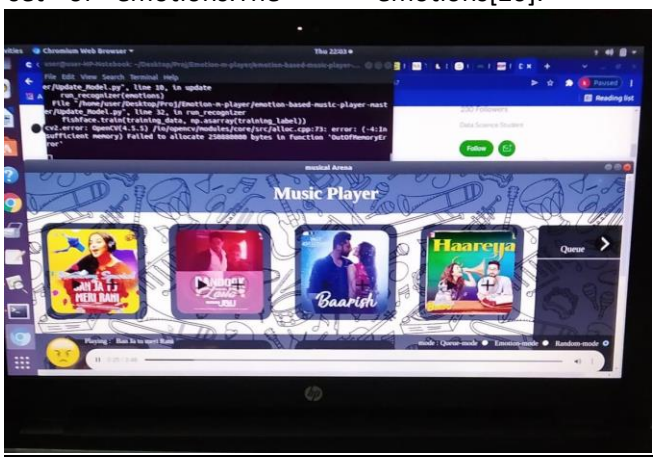

Fig.6. Music selection and emotion recognition

5. Conclusion

In this study, we proposed a models to select music based on facial expressions that indicate emotion. An emotion-based music recommendation system employing facial recognition technology was suggested in this study. Music has the ability to

relieve tension and all types of emotions. The potential for constructing recommendation systems for music based on emotions has recently increased. In order to recognise emotions and play the appropriate music, the recommended system provides a face-based emotion recognition system.

In today's society, a music player with facial recognition technology is very necessary for everyone. This system has been further improved with features that can be upgraded in the future. The mechanism for improving music playback that occurs automatically uses facial expression recognition. The RPI camera's programming interface allows for the detection of facial expression. An alternate approach built on feelings other than revulsion and terror that are not recognised by our system. To assist the automated playing of music, this feeling was introduced.

References

[1]. AnaghaS.Dhavalikar and Dr. R. K. Kulkarni, "Face Detection and Facial Expression Recognition System" 2014 International Conference on Electronics and Communication System (ICECS -2014).

[2]. Yong-Hwan Lee , Woori Han and Youngseop Kim, "Emotional Recognition from Facial Expression Analysis using Bezier Curve Fitting" 2013 16th International Conference on Network-Based Information Systems.

[3]. ArtoLehtiniemi and Jukka Holm, "Using Animated Mood Pictures in Music Recommendation", 2012 16th International Conference on Information Visualisation.

[4]. F. Abdat, C. Maaoui and A. Pruski, "Humancomputer interaction using emotion recognition from facial expression", 2011 UKSim 5th European Symposium on Computer

[5]. T.-H. Wang and J.-J.J. Lien, "Facial Expression Recognition System Based on Rigid and Non-Rigid Motion Separation and 3D PoseEstimation," J. Pattern Recognition, vol. 42, no. 5, pp. 962-977, 2009.

[6] Renuka R. Londhe, Dr. Vrushshen P. Pawar, "Analysis of Facial Expression and Recognition Based On Statistical Approach", International Journal of Soft Computing and Engineering (IJSCE) Volume-2, May 2012.

[7] AnukritiDureha "An Accurate Algorithm for Generating a Music Playlist based on Facial Expressions" : IJCA 2014.

[8] Bruce Ferwerda and Markus Schedl "Enhancing Music Recommender Systems with Personality Information and Emotional States": A Proposal: 2014.

[9] S. Mithen, "The singing neanderthals: The origins of song, verbal, and physique". London, England: Harvard University Press, 2006.

[10] F. Randri, "Emotion-based music reference system by a bottomless reinforcement learning method," *Analytics Vidhya*, 26-Jan-2021. [Online]. Accessible: https://medium.com/analytics-vidhya/emotion-based-music-recommendation-system-using-a-deep-reinforcement-learning-approach-6d23a24d3044.

[11]S. A. Nash, "Charles Darwin on the appearance of human emotions," *Brain World*, 12-Feb-2020. [Online]. Accessible: https://brainworldmagazine.com/charles-darwin-on-the-appearance -of-human-emotions/.

[12] K. Cherry, "The 6 types of simple sentiments and their consequence on human manners," *Verywell Mind*, 03-May-2018. [Online]. Available: https://www.verywellmind.com/an-overview-of-the-types-of-emotions-4163976.

[13]J. van Wyhe, "Darwin, C. R. 1872. The appearance of the sentiments in man and animals. London: John Murray. First edition," *Org.uk*. [Online]. Accessible: http://darwin-online.org.uk/content/contentblock?itemID=F1142&basepage=1&hitpage=1&viewtype=text.

[14]"LDA vs. PCA," *Towards AI*, 26-Jan-2022.

[15] X.-C. Yuan and C.-M. Pun, "Feature mining and native Zernike flashes based symmetrical invariant watermarking," *Multimed. Tools Appl.*, vol. 72, no. 1, pp. 777–799, 2014.

[16]J. K. Nuamah, Y. Seong, and S. Yi, "Electroencephalography (EEG) ordering of intellectual jobs based on task arrangement index," in *2017 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)*, 2017, pp. 1–6.

[17] Gouri Sankar Mishra, Pradeep Kumar Mishra, Parma Nand, Rani Astya, and Amrita, "User Authentication: A Three Level Password Authentication Mechanism", in Journal of Physics: Conference Series 1712,2020, 012005 doi:10.1088/1742-6596/1712/1/012005

[18] Rachna Jain, Abhishek Sharma, Gouri Sankar Mishra, Parma Nand, and SudeshnaChakraborty,

"Named Entity Recognition in English Text", in Journal of Physics: Conference Series 1712, 2020, 012013, doi:10.1088/1742-6596/1712/1/012013
[19] G. S. Mishra, P. Nand and Pooja, 'English text to Indian Sign Language Machine Translation: A Rule Based Approach', International Journal of Innovative Technology and Exploring Engineering (IJITEE), 2019, 8(10)
[20] Gouri Sankar Mishra, K. K. Ravulakolluand A. K. Sahoo, “Word based statistical machine translation from English text to Indian Sign Language”, in Journal of Engineering and Applied Sciences,12(2), 2017, p.481-488.
[21] S. Tyagi, andGouri Sankar Mishra, “POS Tagging Using Support Vector Machines and Neural Classifier”, in International Journal of Computer Science And Technology, Vol. 7, Issue 2, 2016, ISSN : 2229-4333
[22] Pradeep Kumar Mishra, Ali Imam Abidi and Gouri Sankar Mishra, “Improved Methodology For Personality Assessment Using Handwritten Documents”, in Journal of Positive School Psychology, Pages 3263-3273,pub 2022/6
[23] Pradeep Kumar Mishra, Ali Imam Abidi and Gouri Sankar Mishra, “Baseline and its Slant Based Personality Assessment from Handwritten Documents”, in International Journal of Mechanical Engineering PP 2947-2953, pub 2022/5
[24] SamikshaKumari, Karan Kumar Singh, Gouri Sankar Mishra and Parma Nand, "A Comparative Study Of Security Issues And Attacks On Underwater Sensor Network", Accepted, to be published in Lecture Notes in Networks and Systems, 2021, LNEE: https://www.springer.com/series/15179
[25] S. Parhar, A. Roy, K. Kumar, A. Kumar and G. S. Mishra, "Lung Field Segmentation of X-ray Images by Normalized Gradient Gaussian Filter and Snake Segmentation," 2021 2nd International Conference for Emerging Technology (INCET), Belagavi, India, 2021, pp. 1-5, doi: 10.1109/INCET51464.2021.9456146.
[26] Syed Faraz Ali, Gouri Sankar Mishra and Ashok Kumar Sahoo, “Domain bounded English to Indian sign language translation model”, In International Journal of Computer Science and Informatics, 2013, p. 41-45