International Journal of Soft Computing and Engineering (IJSCE)
ISSN: 2231-2307 (Online), Volume-14 Issue-6, January 2025

Human Mood Detection using Image Processing and Machine Learning and Deep Learning

Aloke Lal Barnwal, Rupashri Barik

Abstract: ***This work aims to develop an efficient algorithm to automatic the detection of emotions based on facial expressions. Which involves the use of computer vision & machine learning techniques to classify the emotions of individuals or groups in real-time using an image. As mood detection refers to the process of using various techniques and tools to identify or recognize the emotional state or mood of an individual based on their facial expressions. Its purpose will be to provide insights into the psychological state of individuals for various applications such as mental health diagnosis. It typically involves the use of machine learning algorithms and natural language processing techniques to analyze and interpret human behavior. This approach also uses deep learning models to learn the features of facial expressions and detect the corresponding emotions. The results show that the proposed algorithm is accurately detecting emotions from images with better accuracy & less false detection, which can be suitable for use in various applications such as healthcare, entertainment and social media.***

Keywords: Haar Cascade, Convolution on Image, Deep Face, LBP

I. INTRODUCTION

Families have been steadily getting smaller over the past few decades, frequently switching to nuclear family forms for a variety of reasons. Smaller family sizes may be on the rise voluntarily or as a result of uncontrollable external factors. The idea of blended families, which combine different family members, is rising in popularity. The incidence of both parents working outside the home, which limits family time, is a noticeable development. Family members may feel isolated as a result of this circumstance, which could lead to major emotional health problems, including depression. Furthermore, emotional challenges can stem from different factors, including genetic predisposition, where certain emotional traits are inherited through mutations that vary from person to person. Biological factors, such as neuroimmunology disorders and hormonal imbalances, can also contribute to emotional difficulties. Psychosocial reasons, encompassing social and psychological influences, play a significant role in shaping emotional well-being.

Mood detection is the process of identifying and analyzing an individual's emotional state through various means such as facial expressions. It can be useful in a variety of contexts, including mental health treatment, marketing research, and customer service. There are various methods for detecting mood, including machine learning algorithms, sentiment analysis software, and human observation. Mood detection technology has improved significantly in recent years, with advancements in artificial intelligence and computer vision making it possible to accurately detect and analyze emotions in real-time. However, mood detection is not a perfect science, and there is always a margin of error in any system that attempts to analyze human emotions.

In this context, software applications like the one you're working on hold promise for enhancing human interaction and decision-making abilities. By automating the process of identification and detection of emotional states, the software aims to improve the overall human experience. The development of such software involves a series of crucial techniques. One notable technique is the conversion of images from full-color RGB format to grayscale, or gray level. This conversion eliminates the distractions caused by colors, allowing viewers to focus on other key aspects of the image, such as subject matter, textures, shapes, and patterns.

Another essential technique is the application of convolutions on image layers. The Blurbox filter, for instance, softens the image by averaging the pixel values of each pixel with those of its surrounding pixels, resulting in a smoother appearance. The sharpening filter enhances image clarity and eliminates distortion induced by cameras, directing attention to the central elements. The emboss filter creates a three-dimensional effect that gives the image a raised appearance, simulating the characteristics of an embossed object.

A significant step in the software's development is the use of a Haar cascade classifier. This classifier employs multiple filters to detect various facial components within an image. By identifying these components, it becomes possible to extract crucial information about the emotions conveyed by the subject. The deepface technique is then applied to assess the emotions present in the image. Deepface employs the components detected by the Haar cascade classifier to determine the emotional state of the subject. This final step is pivotal in automating and refining the process of emotional state detection.

In summary, this work focuses on creating software that automates the identification and detection of emotions within images. Through techniques such as grayscale conversion, convolution filters, Haar cascade classifiers, and deepface analysis, the software aims to enhance human interactions and decision-making abilities, ultimately contributing to emotional well-being and improved quality of life.

Manuscript received on 15 July 2023 | First Revised Manuscript received on 27 October 2024 | Second Revised Manuscript received on 05 January 2025 | Manuscript Accepted on 15 January 2025 | Manuscript published on 30 January 2025.

*Correspondence Author(s)
Mr. Aloke Lal Barnwal, JIS College of Engineering, Kalyani (West Bengal), India. Email ID: alok.mis.999@gmail.com, Orcid ID: 0009-0001-7476-3364
Ms. Rupashri Barik*, Department of Information Technology, JIS College of Engineering, Kalyani (West Bengal), India. Email ID: rupashri.barik@jiscollege.ac.in, Orcid ID: 0000-0003-0987-679X

II. BACKGROUND STUDY

In this project concepts were used like image processing which is used for manipulation and analysis of digital images using computer vision. It involves techniques that are used to enhance or modify digital images in various ways such as improving image quality, removing noise, correcting distortions, changing the colors, and extracting important features from the images. It consists of few additional feature like image segmentation -here an image is converted into a collection of regional pixels that are represented by a mask or a labeled image and by dividing the image into smaller segments so that the important parts/segments will only be processed from the whole image here multiple filters were used to calculate feature points by subtracting darker region - brighter region ,feature extraction -it is the processing where the important parts/relevant features from the image is extracted for future analysis and uses i.e. the feature segmented in the previous steps ,convolution layer it consists of 3 important features -blurbox It is a straightforward blur where each pixel is set to the average of the pixels surrounding it. It may be written as a discrete convolution of two functions f[n] and g[n], where f[n] represents the image's discrete pixel values and g[n] is the kernel i.e. to smooth out the image, sharpening is used to sharpen the image (i.e. retrieve any lost feature point else highlight it). There are three main reasons behind image sharpening: To overcome blurring introduced by the camera, to draw attention to particular areas & To increase legibility. Any camera's RAW data is slightly unsharp. So the sharpest transitions are used to average out. It also increases contrast which attract human eye i.e. overall enhancing the image & emboss which forms a 3D mold that stick out from the surface which involves highlighting the edges and other feature points. Then the haar cascade classifier which was originally developed by Paul Viola & Michael Jones [4] to detect faces and object but here the proposed work is only going to detect human faces and to do so they created different kind of filters .Then feature points were calculated by subtracting darker region - brighter region then with the help of feature points and filter facial objects were detected still an image can have multiple feature point which may or may not be needed to detect and an image can have upto 1,80,000 feature points which later on decreased to 5k-6k feature point by using *integral image & Adaboost* then to make it even faster additional cascade was introduced - Here the subset of all 6000 features will again run on the training images to detect if there's a facial feature present or not.

The idea behind this is, not all the features need to run on each and every window. And at last Deepface is used for Mood detection [5] which was originally developed by facebook. And can analyze these expressions to detect various emotions such as happiness, sadness, anger, surprise, and disgust. This is achieved through the use of convolutional neural networks (CNNs), which is trained on large datasets of labeled facial images to recognize patterns associated with specific emotions [1].

The general view by discussing few papers using different technology where authors have used Local binary patterns in short LBP for feature extraction which gave them basically a couple of advantages like

i) low cost while calculation
ii) Resistance to irregular rising and falling in image grayscale values i.e. makes it a bit faster in some cases

It could also detect very small structures/microstructure like edges, spots, flat area by understanding histogram Its working algorithm is also simple like here also initially to convert the image into grayscale then for each and every pixel in the image, need to select the p neighbors that are surrounding the center pixel then it is required to take the center pixel and set it to threshold (i.e. compare the center pixel with surrounding pixels if center is greater than or equal to center pixel update/set its value to 1 else 0) for its p neighbors. Next compute the LBP [2] values sequentially in counterclockwise then take/write a binary number consisting of digits adjacent to the center pixel. This binary number (or its decimal equivalent) is called LBP-central pixel code and, further, is used as a characteristic selected local texture.

Whereas some have used SVM (support vector machine) [3] for detecting faces. As in this paper authors have used multiple calculation for different features extraction like skin color detection using RGB, YCbCr and HLS, edge detection using threshold value and by calculating intensity gradient of the image, extracting face features [7] by calculating luminance and chrominance value the it is trained as SVM [8] classifier on a set of positive and negative examples of faces.

Positive examples are images containing faces, while negative examples are images without faces [11]. The SVM learns to separate these two classes based on the extracted features [12]. Then trained SVM is used to detect new images This is done by sliding a detection window over the image and applying the SVM to the features within the window [13]. If the SVM output is above a certain threshold, a face [10] is detected.

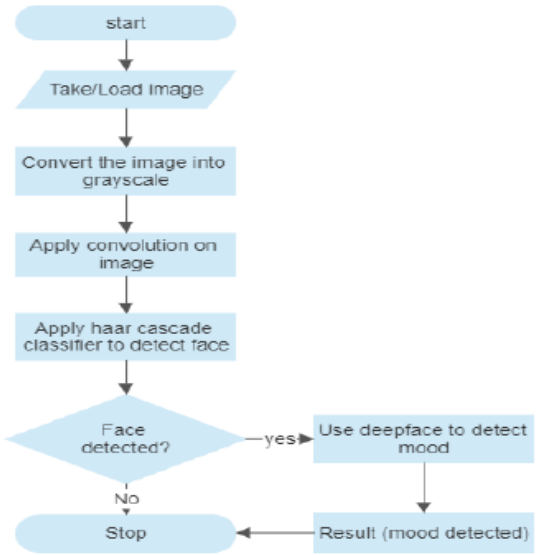
But some authors have used opencv for computer vision i.e. a package / infrastructure to edges and other components of an image then used haar cascade classifier for detecting features like eye, nose, eyebrow, etc. by applying multiple filters on to the image then feature points are calculated and to make it faster integral image & Adaboost is used [14]. And by calculating the distance from one edge to another from features the emotions/mood is detected or predicted [15]. Still some authors have different models that use a convolutional neural network (CNN) to extract features and detect faces in the image [6].

Here the detected faces are then aligned to a standard pose using a geometric transformation, which helps to improve feature extraction and mood detection. It extracts facial features from the aligned faces using a deep neural network. Which include facial landmarks, head pose, and other facial attributes. Then the extracted features are used to determine the mood of each detected face. And after training the model it can classify moods into one of seven categories: neutral, happy, sad, angry, fearful, surprise, and disgust [5].

III. METHODOLOGY

Initially, it is required to install all the necessary packages and import all necessary libraries use OpenCv (computer vision which provide infrastructure

for computer vision applications) then load the image after that convert the image into gray level then apply convolution layer which consists of 3 feature firstly blurbox to smooth out the image secondly sharpening which overall enhances the image and highlight edges and fine details of it lastly Emboss is a technique used to create a 3D-like effect on an image. This technique works by creating a beveled or raised edge on the image, giving it a sense of depth and dimensionality. It involves creating a convolution kernel that highlights the edges of an image. The kernel is then applied to the image, resulting in a new image where the edges are enhanced and appear to be raised or recessed. Then the haar cascade classifier is implemented to detect the face of an image by using filters and calculating feature points then calculate use integral points to make the calculations easier and adaboost to make it faster. Lastly Deepface is implemented which is a Lightweight Face Attribute Analysis Framework and is used in this work to detect the emotion using imp feature point (6k) from previous step as well as it can detect Age, Gender, Emotion, Race. Then a user defined dataset has been taken to train this algorithm.



[Fig.1: System Flow of the Proposed Algorithm]

A. Representation of Proposed Algorithm

Step 1- Take a new image or load it from data set

Step 2- Convert the image into grayscale to remove any distraction of color and help the software to focus on other aspects of the image.

Step 3- Apply convolution on image i.e. subdivided into three steps first use Blurbox to soften the image, secondly use sharpening to focus attention on to the image and remove camera-induced distortion and thirdly used Emboss to give it a 3D mold that protrudes from the surface after Application features will be clearly raised in high relief.

Step 4- Apply haar cascade classifier to detect face using different filter onto the Image in order to detect facial parts like eyes, nose, lips, etc. then integral points will be calculated from feature points to make the calculation easier and faster than adaboost and additional cascade [9] will be used to ignore unimportant/unnecessary feature points to make it better and faster. If the face is not detected then it will stop, else it will move to the next step.

Step 5- Apply deepface which will be used for mood detection. This system utilizes a deep neural network to map the facial features of an individual and identify them with high accuracy.

Step 6- Result will be displayed where the mood will be detected from that particular image.

IV. RESULTS AND DISCUSSIONS

After studying multiple approaches, different ways to get better results can be followed. Here, few additional layers were used like convolution layer on the image while preprocessing (i.e. feature extraction stage of image). It (convolution layer) consists of 3 major components like-BlurBox, sharpening, & emboss which explained in detail in methodology/working section. And applying it is observed that the proposed algorithm has less false detection then the previous version if the image is having multiple faces.

Original Image

Output with Earlier Approach

Output with Proposed Algorithm

It can be observed that in original image there are 8 faces but the previous generation is showing 33 faces i.e. 33-8=25 false detections whereas the proposed work is showing 15

faces i.e. 15-8 = 7 false detections so by this output it can be determined that the proposed working model is a little bit better in terms of accuracy. The aim of the work is to provide an improver one of the steps by preprocessing for future reference expecting the overall result will have a great impact on the system.

V. CONCLUSION

Mood detection is a fascinating and rapidly growing field that has the potential to greatly benefit individuals and society as a whole. With the advancements in technology, machine learning algorithms, and the availability of vast amounts of data, it is becoming increasingly possible to accurately detect and analyze an individual's mood in real-time. Mood detection has numerous applications, including mental health diagnosis and treatment, personalized marketing, and improving human-robot interactions. It can also help identify potential risks such as depression, anxiety, and other mental health disorders, allowing for early interventions and treatments. However, there are also potential concerns around the accuracy and ethical implications of mood detection technology, particularly in the areas of privacy and security. As such, it is important for researchers and developers to approach the development and deployment of mood detection technology with care and consideration. Overall, mood detection technology has the potential to be a valuable tool for understanding and improving human well-being, but it is important to approach it with caution and a critical eye towards its benefits and potential drawbacks.

DECLARATION STATEMENT

After aggregating input from all authors, I must verify the accuracy of the following information as the article's author.

- **Conflicts of Interest/ Competing Interests:** Based on my understanding, this article has no conflicts of interest.
- **Funding Support:** This article has not been sponsored or funded by any organization or agency. The independence of this research is a crucial factor in affirming its impartiality, as it has been conducted without any external sway.
- **Ethical Approval and Consent to Participate:** The data provided in this article is exempt from the requirement for ethical approval or participant consent.
- **Data Access Statement and Material Availability:** The adequate resources of this article are publicly accessible.
- **Authors Contributions:** The authorship of this article is contributed equally to all participating individuals.

REFERENCE

1. P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, Kauai, HI, USA, 2001, pp. I-I, DOI: https://doi.org/10.1109/CVPR.2001.990517
2. Caifeng S, Gong S, McOwan P, Facial expression recognition based on Local Binary Patterns, Image and Vision Computing, Volume 27, Issue 6, 4 May 2009, Pages 803-816, DOI: https://doi.org/10.1016/j.imavis.2008.08.005
3. Shah, Parin M., "Face Detection from Images Using Support Vector Machine" (2012). Master's Projects. 321. DOI: https://doi.org/10.31979/etd.wg5s-gyqn
4. Y. Seanglidet, B. S. Lee and C. K. Yeo, "Mood prediction from facial video with music "therapy" on a smartphone," 2016 Wireless Telecommunications Symposium (WTS), London, UK, 2016, pp. 1-5, DOI: https://doi.org/10.1109/WTS.2016.7482034
5. Y. Taigman, M. Yang, M. Ranzato and L. Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Verification," 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014, pp. 1701-1708, DOI: https://doi.org/10.1109/CVPR.2014.220
6. O. Barkan, J. Weill, L. Wolf and H. Aronowitz, "Fast High Dimensional Vector Multiplication Face Recognition," 2013 IEEE International Conference on Computer Vision, Sydney, NSW, Australia, 2013, pp. 1960-1967, DOI: https://doi.org/10.1109/ICCV.2013.246
7. Berg, Thomas, and Peter N. Belhumeur. "Tom-vs-Pete Classifiers and Identity-Preserving Alignment for Face Verification." https://pdfs.semanticscholar.org/5f14/a9595b0796ce6e5338f157b763326c1f632f.pdf
8. L. Wolf, The SVM-minus similarity score for video face recognition, CVPR2013, IEEE Explore. DOI: https://doi.org/10.1109/CVPR.2013.452
9. Y. Sun, X. Wang and X. Tang, "Deep Convolutional Network Cascade for Facial Point Detection," 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 2013, pp. 3476-3483, DOI: https://doi.org/10.1109/CVPR.2013.446
10. Phillips, P., Beveridge, J., Draper, B., Givens, G., O'Toole, A., Bolme, D., Dunlop, J., Lui, Y., Sahibzada, H. and Weimer, S. (2011), An Introduction to the Good, the Bad, & the Ugly Face Recognition Challenge Problem, NIST Interagency/Internal Report (NISTIR), National Institute of Standards and Technology, Gaithersburg, MD, [online], DOI: https://doi.org/10.6028/NIST.IR.7758
11. Riyazuddin, Y. Md., Basha, S. M., Reddy, K. K., & Banu, S. N. (2020). Effective Usage of Support Vector Machine in Face Detection. In International Journal of Engineering and Advanced Technology (Vol. 9, Issue 3, pp. 1336–1340). DOI: https://doi.org/10.35940/ijeat.C5406.029320
12. Sharma, P., & Chauhan, N. (2020). Spoofing Face Detection using LBP Descriptor and KNN Classifier in Image Processing. In International Journal of Recent Technology and Engineering (IJRTE) (Vol. 8, Issue 5, pp. 2281–2286). DOI: https://doi.org/10.35940/ijrte.E5781.018520
13. Bhavana, Naveen, V. J., & Kishore, K. K. (2019). Comprehensive Analysis of Machine Learning Algorithms for Face Detection. In International Journal of Innovative Technology and Exploring Engineering (Vol. 8, Issue 10, pp. 2479–2482). DOI: https://doi.org/10.35940/ijitee.J9542.0881019
14. Kumari, J., Patidar, K., Saxena, Mr. G., & Kushwaha, Mr. R. (2021). A Hybrid Enhanced Real-Time Face Recognition Model using Machine Learning Method with Dimension Reduction. In Indian Journal of Artificial Intelligence and Neural Networking (Vol. 1, Issue 3, pp. 12–16). DOI: https://doi.org/10.54105/ijainn.B1027.061321
15. Sharma, Dr. K., & Garg, N. (2021). An Extensive Review on Image Segmentation Techniques. In Indian Journal of Image Processing and Recognition (Vol. 1, Issue 2, pp. 1–5). DOI: https://doi.org/10.54105/ijipr.B1002.061221

AUTHOR'S PROFILE

Mr. Aloke Lal Barnwal, is a student of MCA, JIS College of Engineering, Kalyani, West Bengal. His area of interest is Machine Learning, Deep Learning. He is willing to continue his research work in the domain of Machine Learning, Deep Learning.

Ms. Rupashri Barik, has completed B.Tech in Information Technology and M.Tech in Software Engineering from WBUT and pursuing Ph.D in Computer Science at KIIT University, Odisha. Presently, she is working as Assistant Professor at Dept. Of Information Technology, JIS College of Engineering, Kalyani, West Bengal. She has teaching experience of more than 17 years. Her main areas of interests are Image Processing, Machine Learning.