ISSN 2063-5346

HOG BASED EMOTION RECOGNITION USING FACIAL IMAGE FEATURES

Koteswara Rao Kolli[1], Phaninder Vinay[2,] Ramesh Boddu[3], K. Naresh[4], G. Mallika[5], E. Sravani[6], Bhanu Prasad[7]

[1] Assistant Professor, Department of ECE, Raghu Engineering College (A), Visakhapatnam, India – 531162.

[2] Professor, Department of ECE, Raghu Engineering College (A), Visakhapatnam, India – 531162.

[3]Assistant Professor, Department of ECE, Raghu Engineering College (A), Visakhapatnam, India – 531162.

[4,5,6] Student, Department of ECE, Raghu Engineering College (A), Visakhapatnam, India – 531162.

Article History: Received: 01.02.2023 **Revised:** 07.03.2023 **Accepted:** 10.04.2023

Abstract

Human emotion recognition using image analysis is a challenging task that has gained a lot of attention in recent years due to its potential applications in various fields, including psychology, marketing, and human-computer interaction. In our present study proposes and evaluates an automated integrated facial image analysis system for the identification of different human. emotions such as (Happy(H), Sad(S), Disgust (D), Fear (F), Anger (A), and Surprise (Su)). Emotion recognition, one of the crucial non-verbal means by which this communication occurs, helps identify the mood and state of the person.

Human machine collaboration becomes more natural if communication happens through the non-verbal means, such as emotions. In this approach, images of the human face are analyzed to extract features related to emotions, which are then classified using machine learning algorithms. In this presents an overview of the state-of-the-art techniques for human emotion recognition using image analysis. The paper discusses the different types of features that can be extracted from facial images, such as facial landmarks, expressions, and color, and the various machine learning algorithms that have been used for classification. In our research we evaluated basic emotions such as happiness, sadness, anger, fear, disgust, and surprise.

Keywords: Cascade object detector (COD), Histogram of Gradient (HOG), Linear discriminant analysis (LDA).

Introduction

Human emotion recognition using image analysis is an emerging field of study that aims to detect and interpret human emotions from facial expressions captured in images. With the proliferation of digital cameras and smart phones, images have become a ubiquitous source of data, making it possible to use them to understand and analyze human emotions [2].

It would be beneficial if the machines are able to understand human emotions, enabling communication to take another step forward. There are several ways in which this emotion recognition can happen and primarily the focus is on speech and

facial-based emotion recognition [1]. There are also many techniques to achieve these facial and speech-based emotion recognitions including deep learning and classical machine learning algorithms [4].

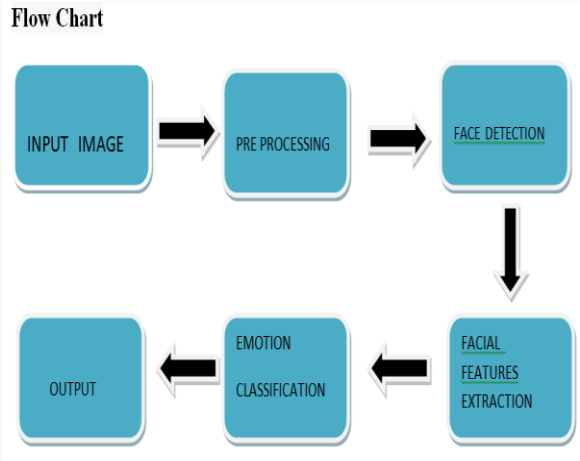
Human emotion recognition using image analysis is to develop algorithms and techniques to automatically detect and classify emotions from facial expressions captured in images. This technology has various applications, including in healthcare, security, and entertainment. Emotion recognition systems typically use machine learning algorithms to analyze facial features such as eye movements, eyebrow position, and mouth shape to identify emotions such as happiness, sadness, anger, and surprise [5].

One approach to human emotion recognition is through the use of machine learning algorithms, such as Histogram of Gradients (HOG) and K-NN, which can be trained on large datasets of labeled facial expressions to recognize patterns associated with different emotions. Other techniques include feature extraction and classification, which involve extracting relevant features from facial images and using them to classify emotions.

In this field has shown promising results, with some studies achieving high accuracy rates in recognizing basic emotions such as happiness, sadness, anger, fear, and surprise. However, challenges still exist, such as recognizing more complex emotions and dealing with individual differences in facial expressions [7].

Human Emotion detection can be obtained by mainly three steps that is pre- processing of image, facial feature extraction in an image and classification.

FLOW CHART



1.pre-processing

pre-processing is an essential step in human emotion detection using image analysis .It can be significantly improve the accuracy and efficiency of the algorithm ,making it more reliable for practical applications. The main process of pre-processing is to prepare the image data in a way that makes algorithm to detect and analysis facial expressions accurately. The following are some common pre-processing steps for human emotion detection.

Image resizing: Resizing the input image to a standard size reduces the amount of data to be processed and it can also help in improving the computational efficiency of the algorithm.

Image cropping: cropping the image to focus only on the face region of interest can improve the accuracy of emotion detection.

Noise Reduction: The noise present in the image can be removed for better results. we have to use filtering techniques to reduce the noise present in the image. There are many filtering techniques to reduce noise in the image such as Gaussian Filters, Median Filters, Laplacian Filter and so on.

In our present research we use median filter to reduce noise in the image. The median filter is a popular technique used in image processing, including human emotion detection from images. The median filter is a non-linear filter that replaces the pixel values with the median value of

neighboring pixels in a window of a defined size.

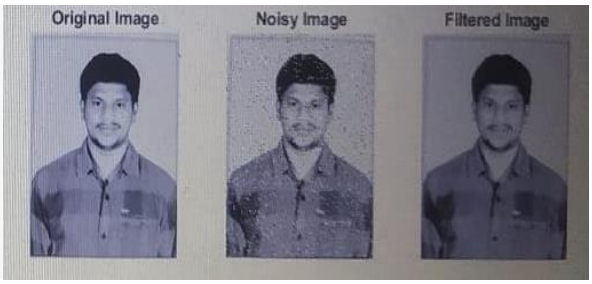


Figure:1 Median filtering technique used to reduce noise present in the image

Facial features extraction:

Firstly, we have to extract the face form the image. The overall face extraction from the image is done first using a Viola-Jones cascade object face detector. It is a popular algorithm to detect face in the image.

Then we have to extract facial features .It is a crucial step involves in human emotion detection. The process involves analyzing the facial features of an individual, such as the eyes, eyebrows, nose, mouth, and chin, to extract information that can be used to determine their emotional state.

We have so many approaches to extract facial features such as such as facial landmark detection and facial expression analysis. Facial landmark detection involves identifying key points on the face, such as the corners of the eyes, nose, and mouth, while facial expression analysis involves analyzing changes in facial features over time to determine emotions.

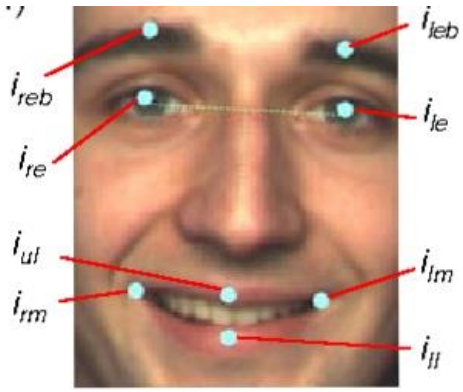


Figure 1.1: Facial Features

Methodology

Local Binary Patterns (LBP): A feature descriptor that detects texture variations in an image, such as wrinkles or creases, which are indicative of certain emotions. The texture information of the face by analyzing the intensity of pixels in the image.

Histogram of Oriented Gradients (HOG)

Histogram of gradients (HOG) is a well-known approach in the field of computer vision that can provide feature descriptions for any given image. HOG is a popular feature extraction technique that captures the local shape and gradient information of an image by analyzing the distribution of gradients in local image patches.

Divide the preprocessed image into small cells, compute the gradient orientation and magnitude for each pixel within each cell, and then create a histogram of gradient orientations within each block of cells. The resulting HOG features capture the texture and shape information of the image at different scales and orientations. It can capture the local shape and gradient information of an image, and it can be used to train machine learning models for emotion classification.

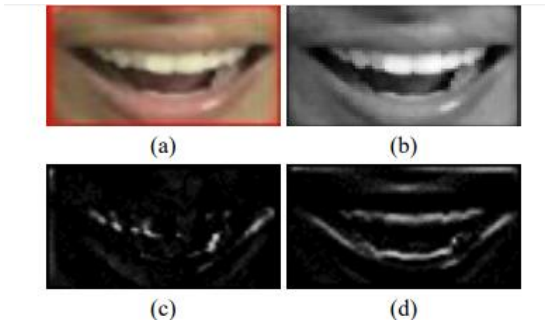


Fig.2. Histogram of Oriented Gradients (a) Detected mouth emotion using Haar Cascade Classifier, (b) Grayscale image (c) X-gradient image (d) Y-gradient image.

A sliding window is moved over a given image and the orientation of each of the pixels is captured within the sliding window. These orientations are also called gradients and all the gradients in a sliding window form a gradient vector. These

gradient vectors are used to form a histogram which reduces the matrix dimensions of the gradient vector significantly. The magnitudes of these gradients that are stored in the form of histogram are then normalized to make them robust to changes in illumination of the image. Figure 2.1 gives a quick glimpse of what HOG does.

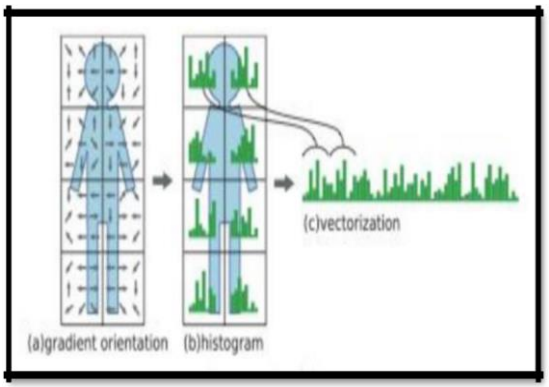


Figure 2.1: Histogram of gradients calculation

K-Nearest Neighborhood

K-Nearest Neighbors (KNN) classification is a machine learning algorithm that can be used for various tasks, including image analysis and classification. K-nearest neighborhood (KNN) classification technique is one of the simplest techniques where the constant "k" is pre-defined based on the dataset [11]. This algorithm considers the similarity of the data points primarily using the distance measures within the dataset which is dependent on the number of neighbors "k" to be considered for a specific data point to be classified. Figure 1.4 gives an overview of KNN classification.

To apply KNN classification to human emotion detection using image analysis, one would first need to train the model on a dataset of labeled images [12]. This dataset would need to include images that have been labeled with the specific emotions they convey (e.g., happy, sad, angry, etc.). Once the model has been trained, it can be used to classify new images based on the emotions they convey.

There are several similarity-based approaches and the most prominently used are the Euclidean, Minkowski and Manhattan distances.

$$Distances : Euclidean = \sqrt{\Sigma}(x - y)\ 2\ ;$$

$$Minkowski = [\Sigma\ [|x - y|]\ n\ ]\ 1\ n\ ;$$

$$Manhattan = \Sigma\ |x - y|$$

Facial expressions database

The facial expression database used in this section has 40 images which is extracted from benchmark KAGGLE (The Japanese Facial Expression) database which originally has 213 images of 10 female models and seven different expressions [11].

Out Of the 40 images, 20 are used for training and other 20 are used for testing. The training images belong to five different female models containing four different emotions (Happiness, Surprise, Sad and Anger) [14]. The testing images belong to eight different female models containing four different emotions (Happiness, Surprise

A) Happy

B) Sad

C) Surprise

D) Anger

Comparison of emotion recognition techniques

Emotion recognition from facial expressions is a popular and efficient way to determine human emotions. Underlying

muscle movements in the human face caused by emotions express variations in facial features which when captured by facial expression identification methods can decode the emotion expressed by the humans.

Various methods and algorithms can be used to determine emotions from facial expressions. One can broadly categorize the methodologies into neural network-based and classical methods-based. Neural networks tend to require more hidden layers which would require intense training through high computing power that becomes suitable to train huge training datasets one time and then later classify emotions for a variety of datasets.

Feature extraction techniques popular in extracting features from facial features are fisher weight maps, principle component analysis (PCA), histogram of gradients (HOG), Wavelets, linear discriminant analysis (LDA), stochastic neighbor embedding, and local features bidirectional.

Facial emotion recognition algorithm:

Step1: Different facial expression images are collected from KAGGLE dataset.

Step2: Convert the color image into grey scale image for accuracy.

Step3: Now remove the noise present in the image by using median filtering technique.

Step4: Then HOG is applied to the image to extract facial features such as eye, mouth, eye brow, nose.

Step 5: Then these extracted features are fed to the classifier for the classification such as KNN classifier.

Step 6: KNN classifier classify the output into different emotions and identify the expression present in the image.

Step 7: Now, the output image is obtained and evaluate the parameters like PSNR, MSE.

Step8: Compare the proposed method images with the existing methods such as SVM, Random Forst.

Parameters

PSNR, MSE parameters are calculated from the below formulas

$MSE = (1/N) * \Sigma i \Sigma j [I(i, j) - J(i,j)]^2$

where N is the total number of pixels in the image,

I (i, j) is the pixel value of the original image at position (i, j),

J (i, j) is the pixel value of the output image at position (i ,j).

It has a range from 0 to infinity, where lower values indicate the better image quality

$PSNR = 10 * \log10(MAX^2/MSE)$

where MAX is the maximum fluctuations in the input image data type. It has a range from 0 db to infinity where higher PSNR values indicat better image quality.

Comparison Table:

Reference	Technique	Accuracy
Shinohara	HLAC + fisher weight maps	68
Lyons	Wavelet + PCA+ LDA	71
Huang M. W	GPLVM+SVM	65.2
Mingwei Huang	SNE+SVM	73
Bin Hua	Bidirectional 2DPCA	79
Proposed method	HOG+KNN classifier	89

Table: Comparison of emotion recognition techniques

Conclusion

In this paper we are proposed a human emotion detection based on the HOG are used to extract the facial features and KNN classifier is used to classify the emotions. It is a promising approach that has shown

good results in recent studies. HOG is a feature extraction technique that is used to detect facial features, while KNN is a machine learning algorithm that is used to classify the emotions expressed in those features. This approach has several advantages, including its simplicity, speed, and high accuracy rates. It also does not require a large dataset for training, which makes it more accessible for researchers and practitioners. However, it is important to note that this approach also has some limitations, including its sensitivity to lighting conditions, image quality, and facial expressions that are not easily detectable by the HOG algorithm.

Overall, human emotion detection using image analysis by HOG and KNN is a promising area of research that has the potential to provide valuable insights into human behavior and improve various applications, such as human-computer interaction, marketing, and healthcare.

References

"[1].A. Mehrabian, "Communication without words", Psychol. Today, (2019)."

"[2].U. Sderstrm and H. Li. Emotion recognition and estimation from tracked lip features."

"[3]. P. Ekman and W. Friesen, "Pictures of Facial Affect. Consulting Psychologist", (2011)."

"[4]. Shbib, Reda & Zhou, Shikun. (2015). Facial Expression Analysis using Active Shape Model. International Journal of Signal Processing, Image Processing and Pattern Recognition.8.9-22. 10.14257/ijsip.2015.8.1.02."

"[5].Widanagamaachchi WN. Facial emotion recognition with a neural network approach. University of Colombo; 2009."

"[6]. Busso, Carlos & Deng, Zhigang & Yildirim, Serdar & Bulut, Murtaza & Min Lee, Chul & Kazemzadeh, AbeLee, Sungbok & Neumann, Ulrich & Narayanan, Shrikanth. (2014). Analysis of emotion recognition using facial expressions, speech and multimodal information. ICMI'04 - Sixth International Conference on Multimodal Interfaces. 205-211. 10.1145/1027933.1027968. "

"[7]. De Silva, L. C., Miyasato, T., and Nakatsu, R. Facial Emotion Recognition Using Multimodal Information. In Proc. IEEE Int. Conf. on Information, Communications and Signal Processing (ICICS'97), Singapore, pp. 397-401, Sept.2010."

"[8].Goren D, Wilson HR. Quantifying facial expression recognition across viewing conditions. Vision Res. 2006 Apr;46(8-9):1253-62. Epub 2005 Dec 20. PubMed PMID:16364393."

"[9]. Michael J. Lyons, Shigeru Akemastu, Miyuki Kamachi, Jiro Gyoba. Coding Facial Expressions with Gabor Wavelets, 3rd IEEE International Conference on Automatic Face and Gesture Recognition, pp. 200-205(1998)."

"[10]. Y.Shinohara and N.Otsu, "Facial Expression Recognition Using Fisher Weight Maps", IEEE Conf. on Automatic Face and Guesture Recognition, pp. 499-504, 2004"

"[11].M.Lyons,J.Budynek, and S.Akamastu, "Automatic Classification of Single Facial Images", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 21, pp. 1357-1362,1999."

"[12]. Huang, M.W.; Wang, Z.W.; Ying, Z.L., "A novel method of facial expression recognition based on GPLVM Plus SVM", ICSP, pp. 916-919, 2010."

"[13].Mingwei Huang; Zhen Wang; Zilu Ying, "Facial expression recognition using Stochastic Neighbor Embedding and SVMs", ICSSE, pp. 671-674, 2011."

[14].Schirmer, A., Chen, C.-B., Ching, A., Tan, L., Hong, R.Y., 2013. vocal emotions influence verbal memory: Neural correlates and interindividual differences. Cogn. Affect.

[15].Behav. Neurosci. 13 (1),80–93. 2. Ko, B.C., 2018. A brief review of facial emotion recognition based on visual information. Sensors 18 (401), 1–20. 3.

[16].Turk, M., Pentland, A., 1991. Eigenfaces for recognition. J. of Cognit. Neurosci. 3 (1),71–86.

[17].Abozaid, A., Haggag, A., Kasban, H., Eltokhy, M., 2018. Multimodal biometric scheme for human authentication technique based on face recognition fusion. Multimedia Tools and Applic 78 (5), 16345–16361.