ISSN: 0976-0172

Journal of Bioscience And Technology
www.jbstonline.com

A Multimodal Approach to Measuring Juvenile Concentration via Facial Expressions and Speech Analysis

Dr.Anbunathan[1]., D.Sharmila[2]

1 Professor, Department of CSE, Malla Reddy College of Engineering for Women., Maisammaguda., Medchal., TS, India

2, B.Tech CSE (20RG1A0505),

Malla Reddy College of Engineering for Women., Maisammaguda., Medchal., TS, India

Article Info

Received: 28-01-2022 *Revised: 15-02-2022* *Accepted: 26-02-2022*

***Abstract*:**

One crucial aspect of interpersonal nonverbal communication is facial expression. The study of computer vision, automation, and artificial intelligence is a rapidly expanding and enduring discipline. Training pictures may be used to identify facial emotions. The emotion identification methods that are now available have several limitations, such as noise. This study presents a Support Vector Machine (SVM) based facial emotion identification technique. Convolutional Neural Networks (CNN) are also used by it for picture training. The kernel trick is the method used by SVM to change the data. Due to its superior performance over other methods already in use, SVM enhances facial expression recognition overall.

Keywords: support vector machines, biometric markers, computer vision, convolutional neural networks, kernel tricks, and machine learning.

1. Introduction

People are sensitive beings. Our emotional states influence not only the most basic processes but also complicated activities and tough choices that we make. Since emotions play a major role in guiding our lives, understanding emotions better helps us understand human behavior in general. It's evident that knowing people's emotional states may be helpful for a variety of purposes, such as improving knowledge of human psychology, studying behavior for better user experiences, creating effective advertising campaigns, and more. The geometric feature-based approach and the appearance feature-based approach are the two primary categories for the many methods to face feature description. The geometric relations between face components are employed in the geometric feature-based method to describe a facial picture. The facial features needed for this, however, must be in precise placements since as a person's look changes, it becomes harder to identify them. As a consequence, this kind of strategy could provide less effective outcomes in certain scenarios. However, the face's overall look is explained by the appearance feature-based method. Facial expression recognition involves three primary phases. The following are the three steps:

1. Face detection: This entails identifying faces in pictures or videos.

2. Facial landmark detection: This technique entails obtaining facial feature information from an identified face. For instance, identifying the contours of the face or characterizing the skin's texture.

ISSN: 0976-0172

Journal of Bioscience And Technology
www.jbstonline.com

3. Analyzing facial feature movement or changes and categorizing information into groups such as smiles, happy or angry emotions, and likes or dislikes for attitudes, facial muscle activation, and emotion categories are the steps involved in facial expression and emotional classification.

An algorithm is used in facial expression recognition to identify emotional states, identify faces, and code facial expressions. It does this by using cameras built into laptops, mobile phones, and desktop computers to analyze faces in pictures or videos. In order to monitor video streams, it can identify various emotions on a person's face as well as business photographs and videos in real time. Applications for facial expression recognition include monitoring stress levels in people, identifying emotional states in mental patients, and creating animated films. It may also be used to identify driver sleepiness. When a driver feels sleepy, smart automobiles may identify it by first sensing its face and then its eyes. It may also be used to identify emotions during an interview to assess whether a candidate's personality matches the position. It is also used in video game testing. Users are requested to play the game for a certain amount of time during the testing process, and their input is used to create the finished product. Biometric indicators are used by facial expression recognition techniques to identify emotions on human faces. The six universal phrases are automatically identified by this method. Nonverbal cues conveyed by facial expression are crucial in interpersonal relationships. These days, a variety of technologies are used to recognize facial expressions. The Neighborhood Edge Directional Pattern (NEDP), implemented in MATLAB, is the current technique. A live picture or video stream will not be produced by NEDP. When using MATLAB, the programming may become sluggish and the visual quality will be subpar. Thus, we are putting out a novel one here, which utilizes Python and is named Support Vector Machine (SVM) in conjunction with Convolutional Neural Network (CNN). SVM is capable of operating on large dimensional data spaces and recognizing face expressions from live images. Python is quicker than other platforms and provides a data structure that is easy to use. As a result, it enhances performance.

2. Proposed Method and Design

The proposed method for facial expression recognition uses the algorithm, Support Vector Machine and neural network, Convolutional Neural Network. The platform used here is Python. Python is a general-purpose, object-oriented, high-level and powerful modern computer programming language. Python uses English keywords frequently where as other languages uses punctuations. It is easy to learn, easy to read, easy to maintain, portable and extendable language. It is processed at run time by the interpreter. You do not need to compile your program before executing it. It can be used to handle big data and perform complex mathematics. The block diagram for proposed method is shown in fig. 1.

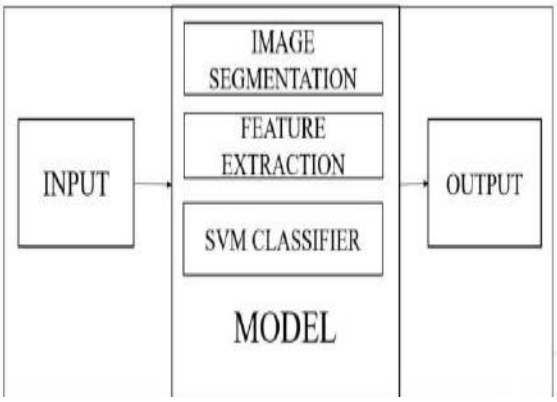

Fig. 1. Block diagram for proposed method

The input is the image. Each image consists of dataset values. The input is obtained by using web cams or it is store it on the OS folder. The image is then sent for segmentation. It involves dividing a visual input into segments to simplify image analysis. Segments represent objects or parts of objects, and comprise sets of pixels. Segmentation is done by using the top-down approach, which is the breaking down of the system into sub-systems. The feature of the segmented image is then extracted. For converting the gradient value to grey scale value, a method called average method is used. The expression for average method is $(R+G+B)3$ where R, G and B represent pixel components red, green and blue. The extracted feature is then fed to the SVM for classification. The training section is done by the CNN. The image would be trained for more number of times repeatedly. The trained image is then compared with the image we want to test. Then the output is displayed in the OS screen.

A. Support Vector Machine (SVM)

Support Vector Machine (SVM) was first heard in 1992, introduced by Boser, Guyon, and Vapnik in COLT-92. Support vector machines (SVMs) are a set of related supervised learning methods used for classification and regression. They belong to a family of generalized linear classifiers. In another terms, Support Vector Machine (SVM) is a classification and regression prediction tool that uses machine

ISSN: 0976-0172

Journal of Bioscience And Technology
www.jbstonline.com

learning theory to maximize predictive accuracy while automatically avoiding over-fit to the data. Support Vector machines can be defined as systems which use hypothesis space of a linear functions in a high dimensional feature space, trained with a learning algorithm from optimization theory that implements a learning bias derived from statistical learning theory. Support vector machine was initially popular with the Neural Information Processing System (NIPS) community and now is an active part of the machine learning research around the world. SVM becomes famous when, using pixel maps as input; it gives accuracy comparable to sophisticated neural networks with elaborated features in a handwriting recognition task. It is also being used for many applications, such as hand writing analysis, face analysis and so forth, especially for pattern classification and regression based applications. The foundations of Support Vector Machines (SVM) have been developed by Vapnik. The classification of image can be done in linearly and non-linearly separable data. Classification in linearly and nonlinearly separable data is shown in fig. 2 and fig. 3.

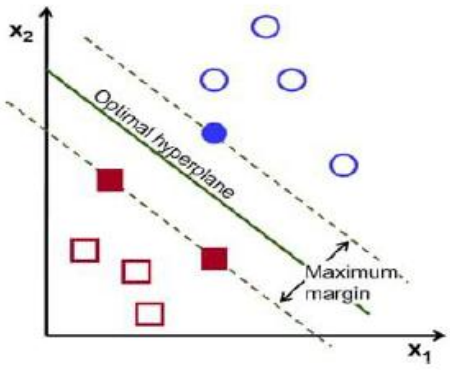

Fig. 2. Classification of linearly separable data

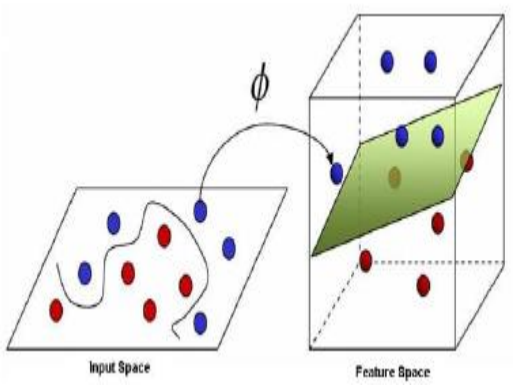

Fig. 3. Classification of non-linearly separable data

Classifying data is a common task in machine learning. Suppose some given data points each belong to one of two classes, and the goal is to decide which class a new data point will be in. In the case of support-vector machines, a data point is viewed as a p-dimensional vector, and we want to know whether we can separate such points with a (p-1)-dimensional hyper-plane. This is called a linear classifier. There are many hyper-planes that might classify the data. One reasonable choice as the best hyper-plane is the one that represents the largest separation, or margin, between the two classes. So we choose the hyper-plane so that the distance from it to the nearest data point on each side is maximized. If such a hyper-plane exists, it is known as the maximum margin hyper-plane and the linear classifier it defines is known as a maximum-margin classifier. However, SVM can be used for classifying a nonlinear dataset. This can be done by projecting the dataset into a higher dimension in which it is linearly separable. In machine learning, a trick known as “kernel trick” is used to learn a linear classifier to classify a non-linear dataset. It transforms the linearly inseparable data into a linearly separable one by projecting it into a higher dimension. A kernel function is applied on each data instance to map the original non-linear data points into some higher dimensional space in which they become linearly separable. Using kernel function, the dot product between two vectors are obtained so that every point is made to a high dimensional data space. Kernel functions are mainly of four types- linear, polynomial, radial basis function (rbf) and sigmoid. Here the kernel function used is the radial basis function. The expression for this function is:

$$e^{\gamma|uv|^2}$$

B. Convolutional Neural Network (CNN)

In neural networks, Convolutional neural network (ConvNets or CNNs) is one of the main categories to do images recognition, images training. Objects detections, recognition faces etc., are some of the areas where CNNs are widely used. CNN were inspired by biological process in that the connectivity pattern between neurons resembles the organization of human visual cortex. Computers sees an input image as array of pixels and it depends on the image resolution. Based on the image resolution, it will see h x w x d (h = Height, w = Width, d = Dimension). The basic block diagram of CNN is shown in fig. 4. The CNN consist of three layers: Convolution layer, Pooling Layer and Fully connected layer.

ISSN: 0976-0172

Journal of Bioscience And Technology
www.jbstonline.com

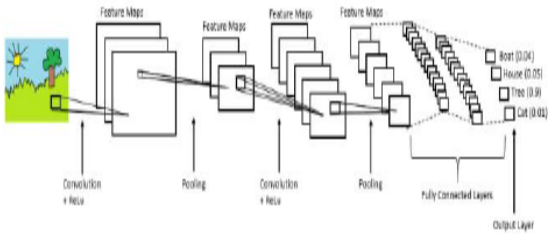

Fig. 4. Block diagram of CNN

3. Result and Discussion

The desired facial expressions can be obtained by using this method. Six universal expressions such as happiness, sadness, anger, fear, disgust, and surprise, can be detected here. Some of the expressions are shown in Fig.5. Multiple faces can be detected in the screen at a time. This means multiple expression can be obtained. The solution for this technique is guaranteed to a large extend. This can be applicable for linearly and non-linearly separable data. For high dimensional data spaces and places where calculations become complex, SVM and CNN can be implemented. The accuracy level of the system is 70- 85%. For obtaining more accurate output, more training of image is required. Hence image training is an important factor in expression recognition. Facial expression recognition also depends on the presence of light. Improper supply of light will also affect its accuracy level. SVM will not be able to classify the image if the choice of kernel function fails. So proper choice of kernel function is important. In this project we have proposed a combination of two methods: SVM and CNN to extract the facial expressions. By using these techniques, we can detect the real time images and it helps to increases the speed of classification. The classified features have good accuracy displaying the expression and facial action units.

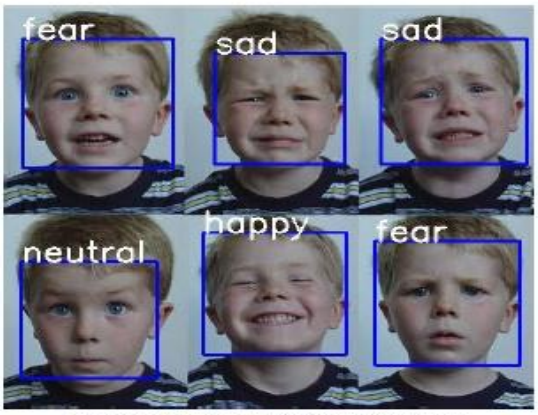

Fig. 5. Detected facial expressions

4. Conclusion

In the area of computer vision, facial activity recognition and facial feature tracking from picture sequences have garnered a lot of interest. One of the most difficult areas of computer vision research to study is computational face expression analysis. Numerous applications, including computer graphic animation, automated face expression identification, and human-computer interaction, need it. Because it can replicate human coding abilities, computer-based facial expression recognition systems, or facial expression recognition, are crucial. Nonverbal communication signals are sent by facial expressions and other gestures, and they are crucial in interpersonal relationships. Two strategies are presented in this research to acquire the facial expressions. Here are six phrases that are universally understood. Several faces may be seen on the screen. SVM and CNN are the two methods we've combined to extract the face emotions and characteristics. where CNN is used for image training and SVM is utilized for picture classification using a kernel approach. By utilizing CNN to train the photos more, we will be able to extract more facial emotions in the future.

References

[1] Caifeng Shan, Shaogang Gong, Peter W. Mc Owan, "Facial expression recognition based on Local Binary Patterns: A comprehensive study," Image and Vision Computing, Volume 27, Issue 6, 2009, Pages 803-816.

[2] Vasanth P.C., Nataraj K.R, "Facial Expression Recognition Using SVM Classifier," Indonesian Journal of Electrical Engineering and Informatics, vol. 3, no. 1.

ISSN: 0976-0172

Journal of Bioscience And Technology
www.jbstonline.com

[3] M. Pantic and M. S. Bartlett "Machine analysis of facial expressions", I-Tech Education and Publishing, 2007.

[4] A. T. Lopes, E. de Aguiar, A. F. De Souza, and T. OliveiraSantos, "Facial expression recognition with convolutional neural networks: coping with few data and the training sample order," Pattern Recognition, vol. 61, pp. 610–628, 2017.

[5] C. Shan, S. Gong, and P. W. McOwan, "Robust facial expression recognition using local binary patterns," in IEEE International Conference on Image Processing 2005, vol. 2. IEEE, 2005, pp. II–370.

[6] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multi class support vector machines," IEEE transactions on Neural Networks, vol. 13, no. 2, pp. 415–425, 2002.

[7] A. Mollahosseini, D. Chan, and M. H. Mahoor, "Going deeper in facial expression recognition using deep neural networks," in Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on. IEEE, 2016, pp. 1–10.

[8] R. C. Gonzalez and R. E. Woods, "Digital image processing," Prentice Hall, 2008.

[9] P. Khorrami, T. Paine, and T. Huang, "Do deep neural networks learn facial action units when doing expression recognition?" in Proceedings of the IEEE International Conference on Computer Vision Workshops, 2015, pp. 19–27.

[10] "Understanding of Convolutional Neural Network (CNN) — Deep Learning" [Online]. Available: https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnndeep-learning-99760835f148