International Journal of Innovative Research in Computer Science & Technology (IJIRCST)
ISSN: 2347-5552, Volume-9, Issue-3, May 2022
https://doi.org/10.55524/ijircst.2021.9.3.8
Article ID IRPV1027, Pages 40-45
www.ijircst.org

Emotion Detection: Comparison of Various Techniques

Jatin Goel[1], Shardul Chauhan[2], and Harshita Jain[3]

[1,3] B. Tech Scholar, Department of Computer Science & Engineering, Inderprastha Engineering College, Ghaziabad, Uttar Pradesh, India
[2] Assistant Professor, Department of Computer Science & Engineering Inderprastha Engineering College, Ghaziabad Uttar Pradesh, India

Correspondence should be addressed to Jatin Goel; goel.jatin24@gmail.com

ABSTRACT- Expressions and body language can tell us a lot about what people are thinking. They are a form of non-verbal communication which tells us about how the person is feeling. It describes the mood of the person like whether he is happy or sad. This detection can be done using various techniques which are already based in the research papers like instrumented sensor technology and computer vision. It means that the expressions can be classified under different techniques like whether motion of the person is still or he is moving. This paper focuses on detecting the emotions of the person using computer vision. Using the Artificial Intelligence Technique and Mediapipe along with Computer Vision we are focusing on various joints in our body and storing their coordinates in a python file created there and then testing our Algorithm to detect the mood of the person. In addition, a dialogue box also pops us while detecting the emotions which tells us about the probability i.e the accuracy of our detection and also tells us about which emotion it is. The current model consists of three emotions, they are happy, sad and victorious i.e Gestures are detected. The algorithm is focusing on the difference between the coordinates and detect the emotions. Detection at a distance might be an issue as the coordinates will be different then. This paper is a thorough general overview of Body Gesture Detection with a brief description of things which are going to be there.

KEYWORDS- Computer Vision, Artificial Intelligence, Gestures, Logistic Regression(lr), Ridge Classifier(rc), Random Forest Classifier(rf), Gradient Boosting Classifier(gb)

I. INTRODUCTION

Communication skills are essential in taking interpersonal interactions forwards. Human simply prefer natural methods of interacting with technologies through the languages that humans have established through time. Aside from words, the indirect natural communication between humans is emotions, which provide a logical way of transmission. It would be advantageous if machines could grasp human emotions, allowing communication to advance further. There are various techniques for emotion classification, with the focus on voice and facial expression identification. Deep learning and traditional machine learning algorithms are among the strategies used to produce these face and speech-based emotion recognitions. Face plays an important part in telling whether a person is happy or sad. Hand gestures are also commonly used when a person has achieved a victory or has been defeated. Using Mediapipe we can map on to the various coordinates (depicted in the image below) and use the coordinates and the angles to detect the emotions and gestures.

Various machine learning techniques have been used for identification of human emotions and hand gestures. In this paper we have shown various machine learning techniques that have been used for emotion detection. These works are discussed one by one in literature survey section. A brief comparison of some of the existing works have been done and shown in the comparison table. Then Mediapipe method is discussed in the methodology used section which denotes how Mediapipe can be used for emotion detection. Accuracy obtained by using some classifiers have also been shown which indicates the accuracy of the Mediapipe technique. Detailed summary of the result obtained is also shown. Finally, the conclusion states the importance of emotion detection and how emotion detection process can be enhanced.

II. OBJECTIVE

The main objective of this paper is to provide a comparative idea between other technologies and this technology which is based on computer vision techniques with regard to gesture and emotion detection. The current paper discusses about emotions like happy, sad and victorious. More emotions can easily be accommodated as a file is created which stores various coordinates according to the emotion mentioned and the coordinates captured by the webcam. It also tells us the probability of correction of detection along with the emotion name.

III. LITERATURE REVIEW

There is a slight difference in Posture and Gesture. Posture focuses on the specific shape whereas gesture focuses on hand movements. Research based on Posture and Gesture can be categorized into two types that are camera-vision based approach and sensor-based approach.

There is a wide field of researches which can be done on the basis of Gestures. Both gestures and emotions try to facilitate communications and also explains whether the other person is interested in continuing the conversation or not. Earlier the gestures were recognized by using sensors and emotions were recognized by various methods like calculating pulse, checking BP, etc. and then the data which

was collected was evaluated using the computers. But, now using Computer Vision it is easier for us to determine Gestures and Emotions using the Web Cam and using various applications of Computer Vision. Mediapipe is a Framework for building machine learning pipelines for processing time-series data like video, audio, etc. Using the coordinates provided by Mediapipe it is very easy to implement gesture recognition.

There are a lot of researches done in the past years. A study by Munir Oudah[1] in 2020 in which he emphasized hand gesture recognition using computer vision. Another study by Myoungseok Yu[2] in 2020 which was a sensor-based model in which real time hand gesture was recognized using CW-Radar. JC Nunez[3] in 2018 provided a research based on Convolutional Neural Networks and Long Short-Term Memory for skeleton-based human activity and hand gesture recognition with comparison between different platforms of GPU and CPU. A study by RF Pinto[4] in 2019 based upon static hand gesture recognition using convolutional neural networks. The procedure in it involves the application of morphological filters, contour generation, polygonal approximation, and segmentation during preprocessing, in which they contribute to a better feature extraction. Finally, Sria[5] published a research paper in 2019 based upon hand- gesture recognition using two-antenna doppler radar with deep convolutional neural networks. However, none of the review paper tells us about the camera type, distance and probability of correct recognition.

These techniques have also provided good outcomes but cannot be implemented on all people. In case there is an elderly people who has various medical issues and recognizing the gestures and emotions using sensors won't be comforting him. Sensors should be used very carefully as they have adverse effects like skin burn, infection(when same sensor is used on various people), also using sensors are also expensive. Some of these problems were addressed in a study by Lamberti and Camastra who developed a computer system based on colored marked gloves. Although this study did not require attachment of sensors, it still required colored gloves to be worn.

These drawbacks helped in discovering of new techniques which did not require the use of sensors and not does not require wearing any gloves. It is known as using the inbuilt sensors which are already present in camera. Along with computer vision many open-source software libraries were also found like tensor flow, Mediapipe which made it easier to detect gestures and emotions. These can be used under various applications like robot control, home automation, gaming, etc. These techniques are a prove that earlier used technique of sensor should be replaced by computer vision. When we go deep into this detection, we get to know that there are different cameras for different purposes like night vision cameras for using in night, RGB (Red, Green, Blue) camera, etc.

Figure 1 shows the coordinates in our body which will be used in Mediapipe for proper identification of human emotions.

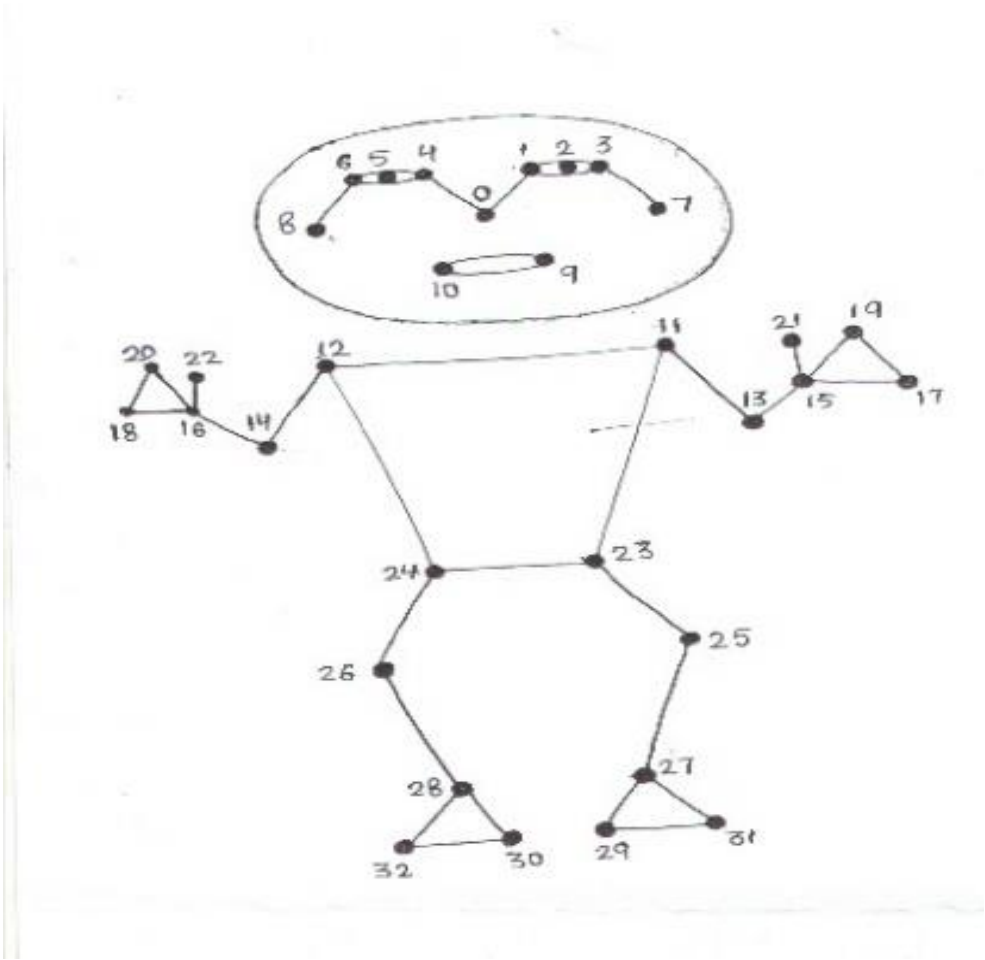

Figure 1: Marking of various coordinates in our body which MediaPipe uses

Figure 2: Coordinates of our body from 0 to 16.

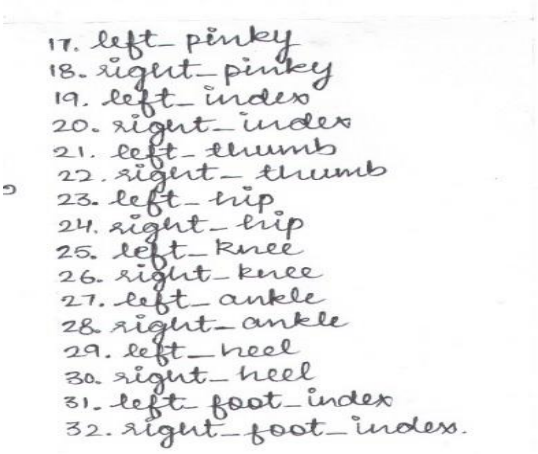

Figure 3: Coordinates of our body from 17 to 32.

Figure 2 and 3 shows the names of all the points which can be used for detection of emotions.

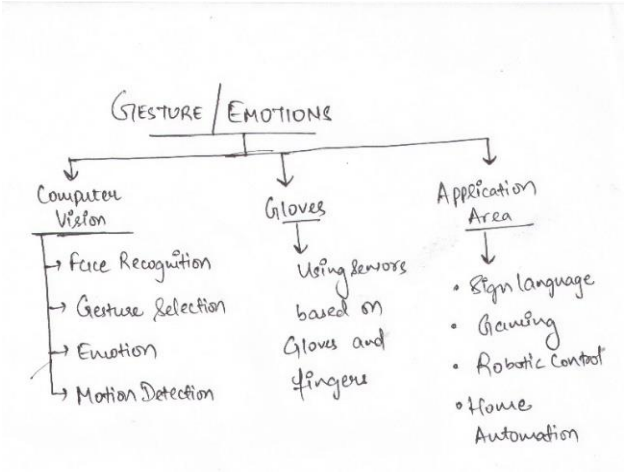

Figure 4: Various ways of detecting gestures.

i) Happy- It is the emotion which is at the top of coords.csv file containing the coordinates. To get the best result neck should be upright and teeth should be visible.

ii) Sad- It is the emotion which is in the middle of the coords.csv file. To get the best result the neck should be downwards.

iii) Victorious- It is the gesture which is at the tail part of the coords.csv file. To get the best result the face should be looking into the camera with the happy emotion and the hands should be pointing up.

Table 1: Research Papers which have used Skin Color Detection for hand gesture

Author	Classify Algorithm	Recognition Rate	No. of Gestures	Distance from Camera
[6]	Maximum distance of centroid fingers	70% to 100%	14 gestures	150 to 200 mm
[7]	Feature matching(Euclidian distance)	90.19%	26 static gestures	100mm
[8]	Distance transform method & circular profiling	100% > according limitation	6 gestures	-
[9]	Contour matching difference with the previous	-	Hand segment	-
[10]	(SPM)classification technique	98.75%	Hand segment	-
[11]	Euclidian distance	82.67%	15 gestures	-
[12]	Combine information from multiple cures of the motion, color and shape	100%	5 hand gestures	-
[13]	Haar- like directional patterns & motion history image	93.13 static static 93.07 dynamic percept	2 static 4 dynamic gestures	(<1) mm (1000-1500)mm(1500-2000)mm
[14]	2D haarlets	99.54%	Hand gestures	1000mm

Table 1 shows the accuracy of some of the exiting techniques that have been used for detection of skin for identification of hand gesture.

IV. METHODOLOGY

With the use of Mediapipe library we can access various inbuilt libraries in it. Mediapipe is used in recognition of various things like face detection, pose detection, gesture detection, etc. with the help of various programming languages like android, IOS, Python, C++and JS. In this project we have used mediapipe with python to detect gestures and emotions of the person. We have used four libraries namely mediapipe which provides face and pose models, opencv which is used for image processing, pandas which is used to work with tabular data as dataframes and scikit-learn to build custom ML-Models. Video Capture algorithm is used to access the webcam in which other working is also done. A python file named coords.csv is also created in which coordinates for various gestures are stored and using which we are detecting the gesture and emotions. Four classifiers are mentioned in this project namely Logistic Regression, Ridge Classifier, Random Forest Classifier, Gradient Boosting Classifier. We have used Random Forest Classifier to classify between the various emotions. 'q' is our ord key which is used to terminate the window which we created using capture function.

Table 2: Research Papers that used appearance-based gesture detection

Author	Classify algorithm	Recognition Rate	No. of Gestures	Dataset Used	Distance from Camera
[15]	parallel cascade structure	above 90%	4 hand postures	Positive and negative hand sample collected by author	–
[16]	feed-forward back propagation neural network	92.33%	26 static signs	Dataset created by author	different distances
[17]	palm–finger configuration	93%	6 hand gestures	–	–
[18]	Fourier-based classification	87.7%	9 hand gestures	ground truth data set collected manually	–
[19]	finger model	–	14 static gestures	The test data are collected from videos captured by web-camera	≤ 500 mm

Table 2 shows some of the works related to appearance-based gesture detection. Here each technique is compared om the basis of recognition rate, dataset used, distance from the camera and classification algorithm.

Table 3: Classifiers used

Classifiers	Accuracy
Logistic Regression()	0.9894
Ridge Classifier()	1.0
Random Forest Classifier()	0.9964
Gradient Boosting Classifier()	0.985915

Table 3 highlights use of various classifiers un der Mediapipe. These classifiers are used to identify the recognition rate.

V. RESULT

In this project we have used four classifiers which are mentioned in the table below. These are telling the accuracy of the algorithms and using these classifiers we have detected the gestures with their correct probability, and they are mentioned with proper recognition rate in table 4.

Table 4: Result Obtained

INPUTS		RESULT OF CLASSIFICATION						RECOGNIZED	CORRECTION PROBABILITY
	CLASS OF GESTURES	Happy	Happy	Sad	Victorious	Victorious	Sad		
	1.	0.97	1	0.88	0.90	0.86	0.85	YES	0.91
	2.	0.93	0.94	0.86	0.84	0.92	1	YES	0.915
	3.	0.92	0.92	0.88	0.86	0.90	1	YES	0.913
	4.	1	0.96	0.92	0.92	0.86	0.86	YES	0.92
	5.	1	0.90	0.96	0.90	0.84	0.86	YES	0.91
	6.	0.90	1	1	0.84	0.86	0.88	YES	0.913

With the above result obtained we can say that our model is 90% accurate.

VI. CONCLUSION

Gestures are a way which tells us how a conversation is going on and whether to continue it or not. Guessing them using computer vision is way easier than using sensors as it is very cheap and easily maintainable as compared with sensors because sensors are required to be maintained, is costly and can harm someone who is of old age or has heart diseases. It takes the same effort to build a computer vision algorithm as it takes to build a sensor. The algorithm has a certain characteristic property which helps us to encounter the common issues and achievable a reliable result. These techniques mentioned above has its pros and cons and can perform well in some cases but not in all.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

ACKNOWLEDGEMENT

The authors would like to thank the staff in Inderprastha Engineering College, Ghaziabad, Abdul Kalam Technological University, Uttar Pradesh, India and the participants to conduct the experiment. No external funding was provided.

REFERENCES

[1] Munir Oudah ; Ali Al-Naji ; Javaan Chahl: Gave a review of techniques based on hand gesture recognition in 2020.
[2] Myoungseok Yu: A sensor-based model in which real time hand gesture was recognized using CW-radar in 2020.
[3] JC Nunez: Research Based on CNN, Short-term memory of skeletal based human activity and hand gestures with various comparisons in 2018.
[4] RF Pinto: A research based upon static hand gesture recognition using CNN in 2019.
[5] S Skaria: A research based upon hand-gesture recognition using two-antenna doppler radar with deep CNN in 2019.
[6] Perimal, M.; Basah, S.N.; Safar, M.J.A.; Yazid, H. Hand-Gesture Recognition-Algorithm based on Finger Counting in 2018.
[7] Pansare, J.R.; Gawande, S.H.; Ingle, M. Real-time static hand gesture recognition for American Sign Language (ASL) in complex background in 2012.
[8] Rajesh, R.J.; Nagarjunan, D.; Arunachalam, R.M.; Aarthi, R. Distance Transform Based Hand Gestures Recognition for PowerPoint Presentation Navigation in 2012.
[9] Choudhury, A.; Talukdar, A.K.; Sarma, K.K. A novel hand segmentation method for multiple-hand gesture recognition system under complex background. In Proceedings of the 2014 International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India.
[10] Stergiopoulou, E.; Sgouropoulos, K.; Nikolaou, N.; Papamarkos, N.; Mitianoudis, N. Real time hand detection in a complex background in 2014.
[11] Khandade, S.L.; Khot, S.T. MATLAB based gesture recognition. In Proceedings of the 2016 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India in 2016.
[12] Zeng, J.; Sun, Y.; Wang, F. A natural hand gesture system for intelligent human-computer interaction and medical assistance in 2012.
[13] Hsieh, C.-C.; Liou, D.-H.; Lee, D. A real time hand gesture recognition system using motion history image in 2010.
[14] Van den Bergh, M.; Koller-Meier, E.; Bosché, F.; Van Gool, L. Haarlet-based hand gesture recognition for 3D interaction in 2009.
[15] Chen, Q.; Georganas, N.D.; Petriu, E.M. Real-time vision-based hand gesture recognition using haar-like features in 2007.
[16] Kulkarni, V.S.; Lokhande, S.D. Appearance based recognition of american sign language using gesture segmentation in 2010.
[17] Fang, Y.; Wang, K.; Cheng, J.; Lu, H. A real-time hand gesture recognition method in 2007.
[18] Licsár, A.; Szirányi, T. User-adaptive hand gesture recognition system with interactive training in 2005.
[19] Zhou, Y.; Jiang, G.; Lin, Y. A novel finger and hand pose estimation technique for real-time hand gesture recognition. Pattern Recognition in 2016.

ABOUT THE AUTHORS

Jatin Goel is a student of Inderprastha Engineering College in B. Tech Third year student and has done conceptualization, funding, writing original draft and editing the draft in this project.

Shardul Chauhan is Assistant Professor in Inderprastha Engineering College and has done investigation and supervision of this project.

Harshita Jain is a student of Inderprastha Engineering College in 3rd year and done conceptualization, funding, writing the original draft and editing the draft in this project. She has been the project administrator of this project.