Emotional Intelligence by Face Recognition Using Machine Learning

***Anish Singh, Harshit Gupta, Piyush Rawat, Siddhant Negi**

*Assistant Professor, Department of Computer Science and Engineering, Inderprastha Engineering College, Ghaziabad, 201010,UP, India

Student, Department of Computer Science and Engineering, Inderprastha Engineering College, Ghaziabad, 201010, UP, India

ABSTRACT :-

In this paper, we have developed a face recognition system that uses machine learning, which will be able to detect the effects of recorded or live video and video. In our daily life, we go through different stages and develop a sense of humor. These feelings and emotions are expressed as a facial expression. Business communities today prefer to use emotion marketing. In the emotional market, they try to stimulate customers' emotions to buy products or services. A high amount of research activity does best in controlled data sets (i.e., small data sets with fewer features), while failing to function properly and still challenging data sets varies imagery even on small faces. In recent years, many activities have introduced a final word recognition system, using in-depth reading models. Although emotional recognition is a major undertaking, there seems to be still room for improvement.

KEYWORDS: Emotion detection , Face recognition, Facial Expression.

I. INTRODUCTION

“True value is inevitably manifested by facial expressions, as the opposite will certainly be clearly defined there. The human face is a natural tablet, the truth is written in it. I live in another's face, as I feel living in my own home. ”~ Maurice Merleau.

"Artificial intelligence is growing rapidly, as are robots with faces that can stimulate emotion and make your mirror neurons vibrate."
~ Diane Ackerman.

Artificial Intelligence is a computer science department responsible for developing machines that can perform tasks that require human ingenuity such as speech recognition, language translation, and decision making. Artificial Intelligence is said to mimic human intelligence in machines designed to think and perform human-like tasks. when combined with the positive personality of the human face and emotional intelligence, it opens the door to endless opportunities.

Face recognition has become a popular research topic recently due to increased demand for security and rapid development of mobile devices. There are many applications where face recognition can be used such as access control, identity verification, security systems, monitoring systems, and social networking networks.

II. PROBLEM

In the modern era of everything being digital, we are developing an applications for supporting the digitalization by creating a face emotional recognition project which will help in detecting the emotions of the individual and can be used to check the mental health of patients, analyzing the feedback of various games through emotion detection . in addition to this intensity of emotion and predicting the overall mood

III. METHODS

A. Two-Channel Emotion Recognition Model

This section will explain in detail the proposed emotional monitoring channel with two channels from the ROI phase and the Gabor feature of the first channel, as well as an effective second channel channel network monitoring

.

B. Division of ROI Area

Enable Pel and mean left and right eye positions, Pn mean nose tip points, Pml and Pmr mean key left and right corner angles, X1 and X2 mean horizontal coordinates of left and right facial contours, and Y1 and Y2 show the vertical position of the upper and lower edges, respectively, to connect. Assuming that the position of the eyebrow area is calculated as an example, only the length and width of the area need to be calculated. The eyebrow length calculator is as follows:

$$\mathrm{H_{eye}} = \begin{cases} \frac{|\mathrm{P_{el}:y - P_n:y}|}{2} + \left|\mathrm{Y_2 - P_{er} : y}\right|, \mathrm{P_{el} : y \leq P_{er} : y} \\ \left|\mathrm{Y_2 - P_{el} : y}\right| + \frac{|\mathrm{P_{er}:y - P_n:y}|}{2}, \mathrm{P_{el} : y \geq P_{er} : y} \end{cases}$$

we use *Weye* to represent the width of the eyebrow area. To make the retrieved eye area not only the eye, but also the part of the information in the corner of the eye, the *Weye* calculation directly takes the distance between the left and right edges. The Calculation is as follows:

$$W_{eye} = |X_2 - X_1|$$

Similarly, the calculation of the shape of the mouth in the face image is as follows, using *Hmouth* and *Wmouth* to represent the length and width of the region, respectively.

$$\mathrm{H_{mouth}} = \begin{cases} \frac{|\mathrm{P_n:y - P_{ml}:y}|}{2} + \left|\mathrm{P_{mr} : y - Y_1}\right|, \mathrm{P_{ml} : y \leq P_{mr} : y} \\ \left|\mathrm{P_{ml} : y - Y_1}\right| + \frac{|\mathrm{P_n:y - P_{mr}:y}|}{2}, \mathrm{P_{ml} : y \geq P_{mr} : y} \end{cases} \quad (3)$$

$$W_{\mathrm{mouth}} = |X_2 - X_1| \quad (4)$$

Finally, the rectangular areas of the eyebrows, eye, and mouth are determined using the above calculation method. Interference with non-essential facial features can be technically reduced, and calculation costs can be reduced, by cutting these regions with major facial changes.

C. Gabor Filter

Compared to other waves, the Gabor transform has a unique biological background. The Gabor filter is similar to the frequency and representation of a person's viewing system in terms of frequency and direction, and can extract location information for different frequencies, position shapes and image directions. The special advantage of Gabor filters is their consistency in scales, rotation and rendering. The reason why the Gaboe wavelet is not used for facial recognition is that when speech changes occur, important parts of the face such as the eyes, mouth, and eyebrows will change dramatically due to muscle changes. These components are shown in the image as changes in the gray scale. Mostly, the real and imaginative parts of the wavelet will fluctuate during this time, so the amplitude response of the Gabor filter in these components will be very obvious, so it is very appropriate to exclude local spatial features. In the field of image processing, the two-dimensional Gabor filter is often used for image processing. The kernel function of the Gabor dual wavelet can be labelled as follows:

$$\psi_{\mathrm{uv}}(\mathrm{z}) = \frac{||\mathrm{k_{uv}}||^2}{\sigma^2} \times \mathrm{e}^{\frac{||\mathrm{k_{uv}}||^2||\mathrm{z}||^2}{2\sigma^2}} \times \left(\mathrm{e}^{\mathrm{ik_{uv}z}} - \mathrm{e}^{\frac{\sigma^2}{2}}\right)$$

where u and v represent the direction and frequency of the Gabor wavelet kernel, z = (x, y) represents the pixel position in the image, σ represents the filter bandwidth, and | kuv | 2 / σ2 is used for compensation. in order to reduce the mass spectrum determined by the quantity. The Gabor facial image feature can be obtained by combining a face image with the Gabor wavelet kernel. Assuming that the gray value of the point (x, y) in the face image is set to k, the Gabor feature count is as follows.

$$G_{uv}(x,y) = I(x,y) * \psi_{uv}(x,y)$$

where Guv (x, y) represents the Gabor feature of the extracted image, ψuv (x, y) represents the kernel function of the Gabor wavelet with two sides, and * represents the function of convolution.

D. Feature Fusion

To take full advantage of the key features of the facial expression and create a lack of global representation in the geographical features released by Gabor, features featured by CNN and local features of the ROI region produced by Gabor are feature-integrated. Simply put, the integration of a feature combines many different features released by different algorithms into a new feature with a powerful rating scale in a particular combination. The process of combining the features is shown in Figure 1.

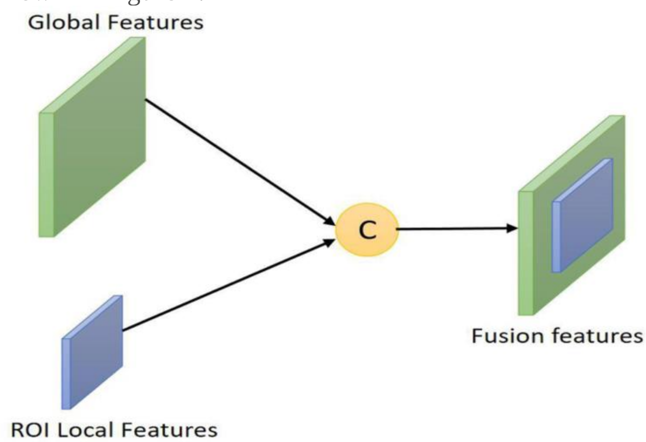

Fig 1. Schematic diagram of feature fusion.

There are three ways to combine features, namely, integration, product and integration. Of the three ways

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 05 | May - 2022 Impact Factor: 7.185 ISSN: 2582-3930

to combine features, the way to combine and integrate is easy to calculate and less computationally expensive. Therefore, this article uses compilation and comparisons to compare the features of Gabor and CNN. Feature integration is done, and local features of the ROI region produced by Gabor and features released by CNN are integrated into a fully integrated CNN layer. Suppose two feature vectors with the same dimension are defined as $X=(x_1,x_2,\cdots,x_n)$ and Y = 1, then the calculation equation for feature stitching is as follows:

$$Z=(x_1,x_2,\cdots,x_n,y_1,y_2,\cdots,y_n)$$

E. Channel Attention Model

To extract additional key features from the facial feature map, I introduced the ultra-lightweight attention module ECA-Net (Wang et al., 2020) to measure the attention of an improved bottle line structure, and to provide greater weight in context. features, What makes the network pay more attention to key features of speech. This structure contains only one parameter k, but the performance improvement that it delivers is very obvious. The main function of this module is to produce the weights of each channel and to learn the correlations between the elements, just as people always ignore important information, but instead focus on information that is useful to us. This module The purpose is to allow the network model to override other non-essential features, increase the emphasis on key features, and the module only adds a very small number of additional parameters.

As shown in Figure 2, the distribution of elements at the size of the area is compressed and subtracted from a two-dimensional matrix to one value, and this value acquires feature information in this space. Then use a fully connected layer to complete the channel size reduction, and then a fully connected layer to complete the channel size, to find the appropriate dependencies between the different channels, generate the weights for each feature channel, and obtain one completely. of installed channels. In order to link between the channels, the essential features of the facial expression channel are made with large weights, and conversely, smaller weights are produced, i.e., the attention span is introduced. It generates channel weights through a one-sided transition in size, and finds the dependence of the correlation between each channel. The side effect of reducing the size of the channel in direct communication between channels and weights is avoided, and obtaining the appropriate dependencies associated with different channels is very effective and accurate in determining the channel's attention. The last two attention modules both duplicate the weights of the generated channels on the first input feature map, and combine the weight-bearing features with the actual features to complete the feature focus in the channel space.

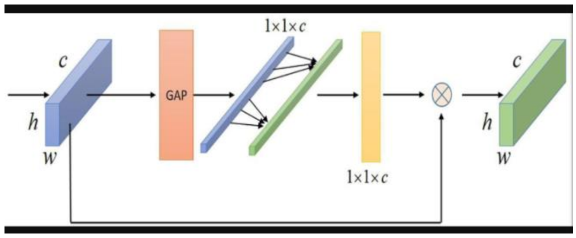

Fig 2. Schematic diagram of the channel attention model.

F. Overall Structure

Fig.3 shows the overall sketch of this algorithm. First, the feature release module is made up of two different CNN branches: CNN's first branch treats the Gabor feature as input. To take full advantage of regional features that have a clear facial expression and rich facial information, the real facial image should be considered in advance and the emotional ROI-related region should be minimized, and then the Gabor wavelet changes should be used to exclude the ROI feature. Since the extracted features are still high in size, a processor analysis of Gabor features is required prior to assembling the feature. This CNN is usually made up of two layers of convolution, which reduce the size of the Gabor features and make it easier to combine the following features. In the second channel, an active channel attention network based on the proposed deep diversification integration is proposed to improve the bottle line structure, reduce network complexity and prevent overcrowding. By designing an efficient attention module, the depth of the feature map is integrated with the location information, focusing more on the removal of key features, and improving the accuracy of sensory perception. Finally, the feature segmentation module separates the components covered by the SoftMax layer.

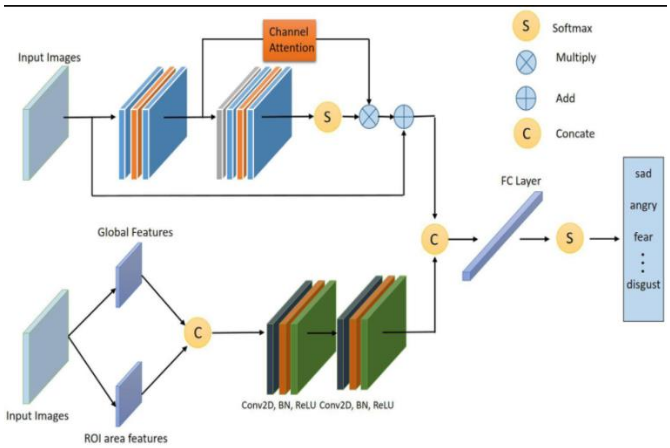

Fig 3. Diagram of the entire framework of the proposed algorithm..

G. Evaluation Method

The overall accuracy rating is used as the test indicator for this study, and its calculation formula is as follows:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}$$

where TP represents the positive predicted models of the model as positive, TN represents the negative predictions predicted by the model as negative, FP represents the negative samples predicted by the model as negative, and FN represents the positive predicted models for the model as negative.

The first step requires computer-assisted grading of the input image on both the x and y directions using the 1x3 and 3x1 detector keys. A horizontal kernel is used in the input image to produce a horizontal gradient image while a vertical kernel is used to produce a gradient image. How the kernels are used does not matter and a straight kernel can be used first to get the same result. We can see an example of a gradient image. As we can see in the gradient image, the edges of the face are retained and can be used for further where TP represents the positive predicted models of the model as positive, TN represents the negative predictions predicted by the model as negative, FP represents the negative samples predicted by the model as negative, and FN represents the positive predicted models for the model as negative.

The first step requires computer-assisted grading of the input image on both the x and y directions using the 1x3 and 3x1 detector keys. A horizontal kernel is used in the input image to produce a horizontal gradient image while a vertical kernel is used to produce a gradient image. How the kernels are used does not matter and a straight kernel can be used first to get the same result. We can see an example of a gradient image. As we can see in the gradient image the edges of the face are retained and can be used for further process.

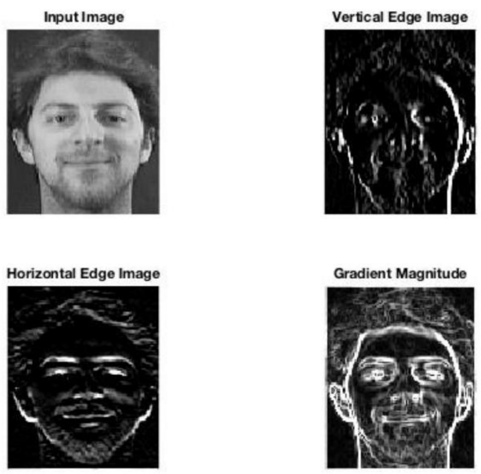

Fig 4. Gradient Image

The next step in processing is to calculate the histogram of the gradient image. Gradient image is divided into a cell grid where each cell is usually 8x8 pixels in size. In each cell, a gradient histogram is calculated. To calculate the histogram gradient two pieces of information are required, size and gradient direction for each pixel . In one cell, the histogram contains nine channels per channel with a range of directions from 0 to 180 degrees. Since we use a range of directions from 0 to 180 and not 0 to 360, these are considered unregistered gradients. Each pixel in the corresponding cell selects a channel based on gradient orientation and that channel's votes based on the gradient size. This process is repeated for each pixel in a cell until the histogram of the gradient vector feature is completed. Now that we have the HOG element vector of each cell, the next part of the process is to get used to it. The purpose of making common vectors is to ensure that the element does not change in any changes in light and contrast. Before we made standard feature vectors, we first created cell blocks in 2x2 groups. Therefore, each block is equal to 16x16 pixels resulting in four vectors of 9x1 HOG element. These four ounces of HOG element can then be combined to form one 36x1 vector element per block. Familiarize yourself with the Input face picture separated by cells. The gradient image of the face is divided into cells. and made this vector of the 36x1 HOG element using the following equation:

$$f = \frac{v}{\sqrt{\|v\|^2 + e}}$$

Fig 5. Image in Cells

where v is a factor of 36x1 HOG vector,

$\| v \|$ the v-shape, and e is a small fixed condition. This routine process is repeated for each block, and the blocks are moved across the entire image so that each cell contributes to more than one HOG feature vectors. Generally, at least the size of the spacing block is desirable. The final step in releasing a feature is to create a

HOG vector feature. This is achieved by incorporating 36x1 standard features in the previous step. So if you have a total of 100 unique block areas each that produces a 36x1 feature vector, the final HOG feature can be length (36) (100) = 3600. Figure 6 below shows a face pattern and a corresponding HOG. feature vector.

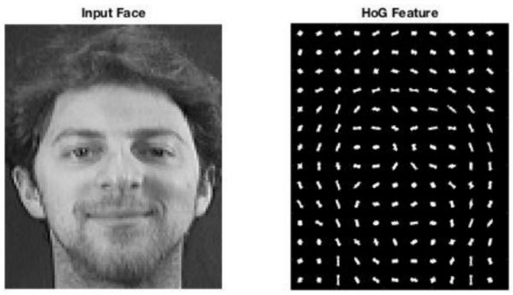

Fig 6. Face Pattern and HOG feature Vector

IV. CONCLUSION

In this paper, we propose a novel feature that incorporates a dual channel expression recognition algorithm based on machine learning theory and emotional philosophy. Because features released using CNN ignore subtle changes in active facial regions, the first method of the proposed algorithm considers the Gabor region ROI feature as input. The active surface region is first separated from the real face image, and the features of this region are extracted using the Gabor transform, with a strong focus on the local area details, in order to make full use of the functional feature data. facial region. To improve bottleneck bottle design, reduce network complexity, and avoid rust, a channel-based network based on deep separation is proposed in the second method. The depth of the feature map is integrated with the knowledge of the area by designing an efficient focus module, focusing more on the removal of key features and improving the accuracy of emotional recognition.

Overall, the revised research in this chapter suggests that emotional intelligence, as reflected in the process of judging facial information, is not universal. It is a multi-faceted skill — not a unique or simple skill — that encompasses a range of structures with intricate interactions with each other.

V. REFERENCES

[1] Bernd Heisele, Purdy Ho, Jane Wu, Tomaso Poggio facial attention: a section based on a global approach. Computer Vision and Image Comprehension. February 2003.

[2] Guodong Guo, Stan Z. Li, Kap Luk Chan Supports vector facial recognition equipment. Image and Vision Computing. January 2001.

[3] Harihara Santosh Dadi, Gopala Krishna Mohan Pillutla Advanced Face Recognition Using HOG Features and SVM Separator. IOSR Journal of Electronics and Communications Engineering, Vol. 11, Issued 4, pages 34-44, July 2016.

[4] Navneet Dalal and Bill Triggs Histograms of Oriented Gradients for Human Detection. Procedures for the 2005 IEEE Computer Organization Conference on computer vision and pattern recognition. 2005. [5] O. Deniz, G. Bueno, J. Salido, F. De La Torre Face recognition using the Histogram of Oriented Gradients. Pattern Recognition Letters. 2011.

[6] P. Jonathon Phillips Vector Vision Equipment Used for Face Recognition. National Institute for Standards and Technology,

[7] Qiang Zhu, Shai Avidan, Mei-Chen Yeh, Kwang-Ting Cheng Rapid Discovery of the Person Using the Cascade of Histograms of Oriented Gradients. Procedures for the 2006 IEEE Computer Organization Conference on Computer Perspective and Patter Recognition. 2006.

[8] Hyeon-Jung Lee, Kwang-Seok Hong A Study on Emotion Recognition Method and Its Application using Face Image , 2017.

[9] Dilbag Singh Human Emotion Recognition System, 2012.

[10] C. Busso, Z. Deng , S. Yildirim , M. Bulut , C.M. Lee, A. Kazemzadeh , S.B. Lee, U. Neumann , S. Narayanan Analysis of Emotion Recognition using Facial Expression , Speech and Multimodal Information, ICMI'04, PY, USA, 2004.

[11] Claude C. Chibelushi, Fabrice Bourel ―Facial Expression Recognition: A Brief Tutorial Overview‖.

[12] Daniel McDuff, Abdelrahman Mahmoud, Mohammad Mavadati, May Amr, Hay Turcot, Rana el Kailouby , "AFFDEX SDK : A Cross-Platform Real-Time Multi-Face Expression Recognition Toolkit", CHI, 2016.

[13] Samira Ebrahimi Kaou, Xavier Bouthillier, Pascal Lamblin, Caglar Gulcehre , Vincent Michalski, Kishore Konda, Sebastien Jean, Pierre Froumenty , Yann Dauphin, Nicolas Boulanger-Lewandowski, Raul Chandias Ferrari, Mehdi Mirze, David Warde-Farley, Aaron Courville, Pascal Vincent, Roland Memisevic, Christopher Pal, Yoshua Bengio, " EmoNets : Multimodal deep learning approaches for emotion recognition in video", Springer, August 2015.

[14] L. Franco, Alessandro Treves ―A Neural Network Facial Expression Recognition System using Unsupervised‖ Local Processing, Cognitive Neuroscience Sector – SISSA.

[15] C. L. Lisetti , D. E. Rumelhart ―Facial Expression Recognition Using a Neural Network‖ In Proceedings of the 11th International FLAIRS Conference, pp. 328—332, 2004.

[16] Ashwani Kumar, Real Time Emotion Detection of Humans Using Mini- Xception Algorithm, IOP Conf. Series: Materials Science and Engineering, INDIA, 2020

[17] K. Sravanthi, G. Jaya Suma, "A Prediction of Emotions for Recognition of Facial Expressions Using Deep Learning", International Journal of Recent Technology and

Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-2S11, September 2019.

[18] Karthick. S, Emotion Detection using Machine Learning, International Journal of Engineering Research & Technology (IJERT), INDIA, 2020.

[19] Syed Aley Fatima, Ashwani Kumar, Ajay Pratap, Syed Saba Raoof, "Object recognition and detection in remote sensing images". 2020 International Conference on Artificial Intelligence and Signal Processing (AISP) 978-1-7281-4458-0/20/$31.00 ©2020 IEEE.

[20] Najmeh Samadiani, , Guangyan Huang , Borui Cai , Wei Luo , Chi-Hung Chi , Yong Xiang and Jing He "A Review on Automatic Facial Expression Recognition Systems Assisted by Multimodal Sensor Data". International Congress of Information and Communication Technology (ICICT 2017).