Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 6, June 2021: 1475-1479

Research On Independent Emotion Recognition From Facial Expression

Shiv Shankar Shukla[*], Avinash Verma[b], Dr.S. Thanga Revathi[c]

[a*,b,c,] SRM institute of science and technology, Kattankulathur, Chengalpattu District, Tamil Nadu-603203

***email:** [a]sr6224@srmist.edu.in,[b] ar1142@srmist.edu.in,[c] thangars@srmist.edu.in

Abstract

Emotions are essential to human communication. It is also helpful to understand the person's attitude and behavior. This study suggests a deeper learning approach to perceiving emotions from the face. We have selected two methods of convolution neural network (CNN) and Artificial neural network (ANN). We will work in all cases depending on the autonomy / dependence. We aim here to see the independent feelings of the subject well. Recent FER focuses on independent emotional recognition and has used the JAFFE dataset (Japanese women's expression) as a benchmark for benchmarking. We have chosen to use voting classifier as our benchmark algorithm and are trying to get better results than in the past. We say the best because it is not the limited data used by the previous paper while it is a combination of learning integration that works with multiple variables and predicts the outcome in terms of high class opportunities. We use the confusion matrix to test the performance of our model..

Keywords: CNN, ANN, vote division, emotions, facial expressions, ReLU layer

1. Introduction

Emotional recognition from the face-to-face results in the field of Human Computer Interaction, Augment Reality, Virtual Reality, Compressed Computer and Advanced Driver Assistance Programs. FER is also active in various areas of sports, health care, education and security especially ATM security. Over the years many efforts have been made to improve communication between humans and computers. Human communication takes place mainly through speech but gestures also play an important role in human interaction. One way to express feelings is to talk face-to-face. The movement or position of the facial muscles determines the appearance of the face. According to some controversial topics, facial expressions convey a sense of emotion to the person who sees them. In fact we can make 10,000 unique facial expressions but there are 7 Universal moods: Happy, Sad, Anger, Disgust, Neutral, Surprise and Fear.

Chibelushi points out that facial expressions have a profound effect on the speaker. has established that the appearance of the speaker's face accounts for about 55 percent of the impact, 38 percent on the opposing voice and 7 percent on the spoken words themselves.

The first section includes a literature review and the second section of this research paper contains information about our model database. We took more than 30000 pictures to train our model with various emotional categories. The third section describes the in-depth study model CNN and ANN and the measurement algorithm. The last paragraph includes concluding remarks.

2 Literature Survey

The facial expression actually tells the story of a person's emotions, many attempts made by an engineer to create a model that can fully detect the Universal emotional state. It is widely used in many fields such as Robotics, Medicine and lie detector. Elham S. Salama [1] has worked on 3D-CNN multi-sensory recognition where he uses a collaborative learning approach. On his proposed route where he combines face details with EEG symbols. Two phases were created, In the first phase two CNN phases were performed to separate the EEG signals. In the second phase a third model is created from the EEG on the basis of fusion mixtures and facial methods. Victor-Emil Neagoe [2] works on in-depth reading to recognize the independent feeling of the subject through facial expressions. This paper proposes a model in which CNN models and deeper beliefs (DBN) were used to train the model and database. This model uses the JAFFE dataset as a benchmark algorithm to train this model. This model is able to detect emotional state. This model has some limitations this model relies heavily on emotional awareness with dependent emotional recognition. This model has produced an emotional perception school through the Deep Belief Network. It was 59.8%. In terms of sensitivity sensitivity based on this model it has excellent sensitivity scores of more than 95 percent. WR Sam Emmanuel [3] has worked on face-to-face recognition in King Saud University magazine. In this paper they have discussed various aspects of facial expressions. In this paper he used the input image and moved on to further development where multiple images were made from a single image and feature extracted and moved to the separation layer where they were able to successfully differentiate the worlds of emotions. Pawel Tarnowski [4] has worked on emotional awareness from the face. This paper presented the result of the state of the seven universal senses on the basis of the face. The features used here were the coefficients that defined the facial feature. They calculate the features of the three-dimensional model. In this paper they used the KNN classification and MLP neural networks in feature classification. [4] Facial expressions play a very important role when it comes to emotional recognition and are used for communication to identify people. The aim of this study was to identify the seven emotional states. WR Sam Emmanuel [5] also worked on the Multi Support Vector Machine (MSVM), and the precision facial recognition algorithm was an algorithm for optimizing whale locusts. Initially, They extract feature vectors in facial images. They used these feature vectors to train the model with a multi-vector support machine.

Livia Graumann [6] who has worked on facial sensory awareness in borderline patients is not affected by major psychological stress. In this paper you have tried to deal with borderline personality disorders and emotional facial expressions to deal with stress. Wafa Mellouk [7] has worked on reviews and insights on facial recognition recognition. In this paper the author focuses mainly on automatic sensory perception and used in-depth learning techniques for spontaneous visual perception. Emad Barsoum [8] worked on training a model to inform facial expressions about the distribution of a crowd label. In this paper the author proposes a system for obtaining crowd data using in-depth reading methods. This proposed model uses the VGG13 network from the beginning of the FER database. In training this model uses Majority Voting, Multilabel Learning, Probabilistic label diagram, Cross-Entropy Loss and a confusion model matrix. This model compares different training schemes. Gathering the truth of the world the discovery of the crowd has been a very influential process. Ninad Mehendale [9] has worked on facial awareness through the convolutional neural network. This paper proposes a two-level flexible neural network. In the first level it removes the layer from the images to avoid the background effect the second level removes the element from the surface and processes it in the partition layer. This paper uses vector expression to maintain the state of all emotions. These two levels of CNN work in sequence and at the final level they adjust weight and exponents. It works better than a single CNN level. This paper trains the model with more than 750K images. Nitya Roopa S [11] worked on sensory perception from the face using an in-depth reading method, This paper suggests a gentle way to find the emotional state. This paper uses a database of emotions collected at Kaggle and used learning transfers to train and initiate the Net v3 model.

Albert ali [12] researched video-based emotional recognition when comparing it to audio-based emotional recognition.

3 Proposed Job

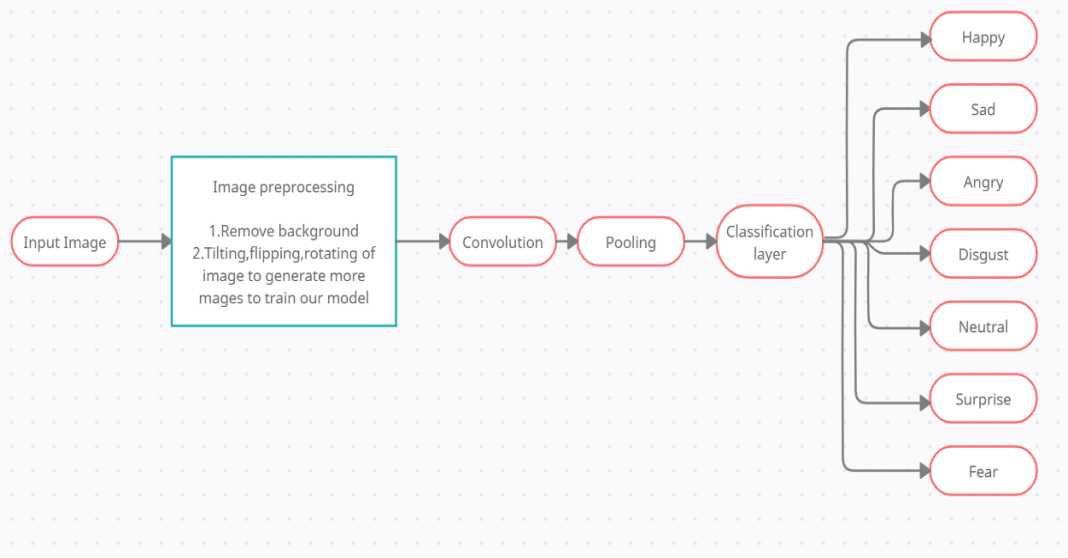


Figure 1. Architecture block diagram.

3.1 Dataset analysis

The Dataset contains images of global feelings of Happy, Anger, Sadness, Disgust, Surprise, Fear and Neutrality. Our training data contains more than 28000 images of various emotional states and verification data contains more than 7000 images of various emotional dimensions. We have taken 80% of the database as the remaining training data as 20% as the validation data. Images in the database are cropped to avoid background influences in training so that our model focuses only on emotion and not behind images.

3.2 Model Training

We follow two approaches - the Convolution Neural Network (CNN) and the Artificial Neural Network (ANN).

1 CNN: Why convolutional neural networks? Knowing that a color image can be represented by three channels and colorless images can be represented by two channels. When transferring an image to the neural network of 200x200x3 channels, it requires a total of 120000 weights and therefore we need a convolutional neural network to avoid such a large number of weights. The Convolution neural network is inspired by the visual cortex.

CNN belongs to the category of deep neural networks, mainly used to deal with images. We can define a convolutional neural network as a standard version of a multilayer perceptron that usually means fully connected networks (used as a separation layer), meaning that each neuron is connected to another neuron in the next layer.

The mathematical function used by the "convolutional neural network" is called convolution which is CNN's first layer. Convolution is a special kind of operation. Four dynamic networks of neural networks use convolution instead of the standard duplication of the Matrix in at least one of their layers.

CNN contains input and output layer, as well as many hidden layers. CNN's hidden layers usually contain a series of convolutional layers that convince duplication or other dot product. The activation function is usually a RELU layer, followed by other combinations such as integration layers, fully integrated layers and configuration layers, called hidden layers because inputs and their effects are hidden by activation function and final agreement.

CNN uses local connectivity and local connectivity pattern between neurons of adjacent layers.

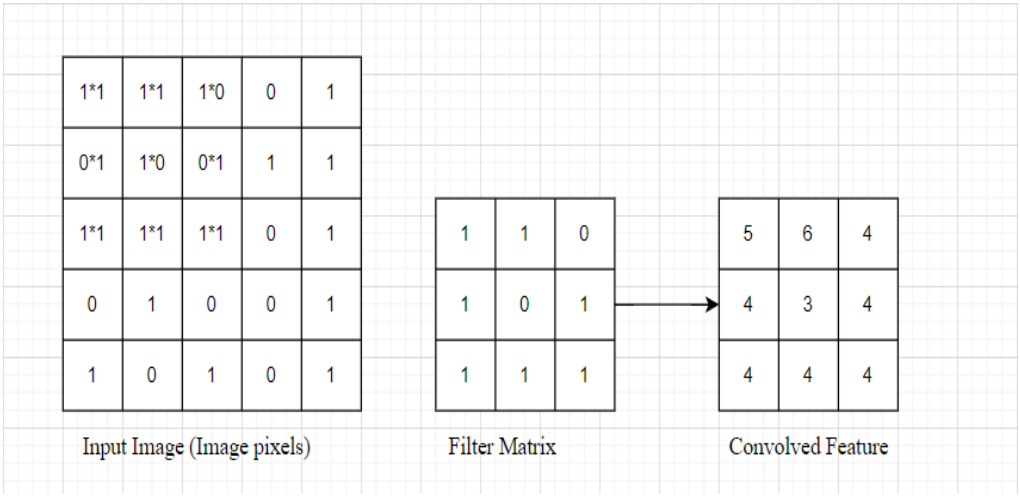
1.1 Convolution: While editing CNN, input position (by number of images) X (image height) X (image width) X (image depth). After passing through the convolutional layer, the image is removed to insert a map, with the shape (image number) X (feature map height) X (feature map width) X (featured channel channels). via the following neural network:

• Variable characters defined by width and height (hyper-parameters).

• Number of input channels and output channels (hyper-parameters).

• The depth of the convolutional layer should be equal to the depth of the convolution feature map.

The conversion layer converts the input and transfers its results to the next layer. Consider the filter matrix



The main purpose of this application is to extract all the features from the input images such as Edge and everything. Convolutional Networks do not need to be limited to one Convolutional Layer. Typically, the first Convolution Layer is responsible for capturing low-level features such as edges, color, gradient shapes, etc. With additional layers, the design conforms to high-quality features and, giving us a healthy network to understand images in the database, is like how we could do it.

1.2 ReLU (Activity function): The ReLU layer is nothing but a function of the convolutional network neural network. This layer or function is used to avoid zero summaries. The ReLU layer called Rectified Layer Unit is defined as Zero where X is negative and contains line dependence where X is reliable.

F (X) = 0, when X<0

X , when X>0

1.3 Pooling: The pooling layer adopts inputs (matrix) from the ReLU layer and reduces that matrix or we can say that the composite layers reduce the input images by considering the important features and ignoring the non-essential features. It uses two concepts to reduce the combination of input images and standard integration. In max pooling the max value is set from the window while calculating the sum of the window.

1.4 Fully Connected Layer (Separation Layer): This is the actual layer of the convolutional neural network where separation occurs. The reduced images from the merging layer are converted to vector here in this layer and matched to the vectors of the subdivisions, a section similar to these vectors is divided into categories.

2 Artificial Neural Network (ANN): Before we talk about ANN let's talk about the neural network because ANN is very similar to a biological neural network. It contains dendrites that take input from other neurons and Nucleus (cell) and axon terminals that contribute to other neuron release. All sensors in this network are connected.

The concept of an artificial neural network (ANN) is very similar to a biological neural network. ANN contains a layer of dendrites in the neural network and nucleus (cell) nodes as well as a layer of Axon terminals in the neural network and synapses equal in weight to the ANN.

As the human body captures input with the eyes (pictures, text), ears (sound), touch with the hand and process it into the brain and train the body to make predictions. When a baby is born he knows nothing about his father, mother and anything. His mind is trained by the environment in the form of images, sound and concept and in the same way he takes things and thinks about them. In the same way our ANN incorporates inputs such as images and audio and video and also trains the model with a labeled database (Database provided with output). It also makes predictions based on training and as the human brain learns by seeing objects, ANN also learns by given input.

3 Voting Classifier: Classifiers are used to classify data objects with a label, Voting partition is part of learning together used to separate data based on the output labels provided. The inclusion of a voting planning process in all classifiers also makes the assumption based on the high probability of model results.

There are two types of voting used - Hard voting: In this type of voting process divides the voting process into all classifiers and makes predictions based on the highest number of votes. Suppose we get the {TTFTFFT} results of certain voters and our voting divider predicts T because of the high probability of a hard vote. In the case of T and F being equal select the category according to the order increase.

Soft voting: In this type of voting, the output is predicted in terms of middle class opportunities. Suppose category A is probable (0.30, 0.47, 0.53) and category B is set for probability (0.20, 0.32, 0.40) so the median probability of category A and B is 0.43 and 0.30 respectively. The winner is therefore category A because we have a high probability of being divided.

Here we will use voting classifier as a benchmark algorithm for our project.

3.3 Predictability and Development

Testing: After completion of training, we will test our model in the test database to see the accuracy of the model.

Predictability: we will be making predictions using models and after gaining the accuracy of the various models, we will be reporting a very good model of emotional recognition.

4 Conclusion

We conclude that by using the models described above we can achieve better performance in our complete model. Our project focuses on predicting the independent feelings of the universe and its dependent emotions. We use an in-depth learning model (Convolutional neural network and neural network) because these algorithms associated with image training and the use of voting planning increase the effectiveness of prediction because it is part of learning together and exploring input into all divisions and select what offers high.

References

[1] Framework of 3D-convolutional neural network with integrated reading methods for multimodal emotional recognition, Egyptian informatics journal, 13 aug 2020, Elham S. Safe Reda A

[2] An in-depth learning approach to independent facial emotion recognition from facial expressions, Department of Information Engineering and Computer Science, University of Trento, Italy.

[3] Research on human face recognition techniques Journal of king Saud University- Computer and information sciences, 5 sep 2018, I.Michael Revina, W.R. Sam Emmanuel.

[4] Face-to-face recognition: Pawel Tarnowski, Marchin kolodziej, Andrzej Majkowski, Remigiusz J. Rak, Warsaw University of Technology, Waesaw, Poland.

[5] Face recognition for optimization based on Multi SVNN and Modified LDP features Journal of visual and image representation 29 april 2019, I.Michael Revina, WR Sam Emmanuel.

[6] Recognition of facial sensations in borderline patients is not affected by major depression, Psychiatry Journal Journal, 16 October 2020, Livia Graumann, Moritz Duesenberg.

[7] Facial emotion recognition using Deep Learning: Review and Understanding, 2020, Wafa Mellouk, Wahida Handouzi.

[8] Training in in-depth face recognition and label distribution networks acquired by the crowd. Emad Borsoum, Cha Zhang.

[9] Facial emotion recognition using the Convolutional Neural Network (FERC). SN used science 2020, Ninad Mehendale.

[10] Facial recognition recognition is based on the cascade of neural networks. New research on multimedia and internet programs. Elzbeita Kukla, Pawel Novak.

[11] Emotional recognition from the face using in-depth learning. International Journal of Advanced Engineering and Technology (IJEAT) August 2019, Nitya Roopa S.

[12] Emotional recognition based on video. Multimodal Wild Behavioral Review 2019, Albert Ali Salah, Furkan.

[13] Automatic recognition of reported and known emotions. Multimodal Wild Behavior Review 2019, Biqiao Zhang and Emily Mower Provost.