e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

LIVE CLASS MONITORING SYSTEM (FACE EMOTION DETECTION)

Ritvesh Mishra*1, Govind Singh Bajni*2, Abhishek Kumar*3, Mohd Sameer*4, Mohd Jawed Khan*5

*1,2,3,4,5Department Of Computer Science & Engineering Greater Noida Of Institute Of Technology, Greater Noida Knowledge Park II, UttarPradesh, India.

DOI : https://www.doi.org/10.56726/IRJMETS39867

ABSTRACT

This paper gives you a current method on facial emotion stages, techniques and datasets. With the rapid increase in online and offline and online classes simultaneously so, there is urgent need to develop technology that can be accurately detect emotions from student expression, which can help in monitoring and engagement of students during the live class on going on. This paper can also be used in healthcare, paper provides proper examination usability of facial emotion and expression detection in both online and offline classes that will enhance the teaching and learning experience. And this paper also deal with issues and challenges comes in detecting emotion and expression with educational contexts and proposes valuable solution to how to deal with those challenges.

Keywords: Facial Emotion, Facial Expressions, Emotion, Detection.

I. INTRODUCTION

The increase popularity of online classes, almost all educational institutions are also adopting the hybrid learning both online and offline classes. This paper can be used in the healthcare sector for detecting the patient emotions. A healthcare framework that incorporatesa facial expression recognition system, which may benefit from the fact that human facial expressions change as one's health changes. A large amount of data is used in various experiments to validate the proposed system. At least 99.95% of the proposed system's trial results show that it can correctly recognize facial expressions.. There are many challenges faced during the online classes and offline, especially in online, so therefore to minimize that challenges it is important to monitor the emotions and expressions of students during the live classes. Emotion detection can help in identifying the students which are losing interest, feel bored and are not motivated, and faculties or teachers to change their method of teaching accordingly to increase their result and productivity [1, 11].

The Facial Action Coding System (FACS) is another way for analyzing facial expressions during the extraction step. Ekman pioneered this strategy. In order to categorizes emotions based on the movements of various facial muscles. Several Action Units (AUs) are used to categorise face expressions, and collections are used to identify emotions. Deep learning is a component of machine learning techniques that may be used for face expression analysis andemotion identification. However, the performance of deep learning is dependent on the amount of the data [4].

According to Feidakis, Degradomes, and Cabela's study, the categorization of emotions basedon logical models, there are 66 emotions that may be grouped into two groups: Anger, Fear, Caution, Happiness, Joy, Love, Sadness, Surprise, and Trust aresome of the primary emotions, and there are 56 secondary emotions. It's extremely hard to evaluate such a large number of emotions, in particular when we need automated recognitionand evaluation [4,5].

The scope of this review is to assess hardware and methods for automatic emotion recognition that can beused in machine learning applications based on as a result of these studies, experimental data is analyzed and automated solutions are generated. This study confirmed the notion of humanizing the Internet of Things as well as Affective Computer Systems. The researchers in this study created systems.[4].

The body of knowledge about teacher emotions and their impacts is expanding at a rapid pace. However, because the study of teachers' emotions is still in its infancy, the questions driving the research were relatively simple and essentially unidirectional. It may be required to investigate the underlying assumption that instructors, in particular, are the source of the correlations between teacher emotions and student results. Because instructors and studentsare both equally a part of the school environment, we conceptualize, express, and conceptualize teacher emotions as part of a system that impacts and is impacted by student outcomes,

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

including students' own emotions, thinking, and behavior's, in this article. [1].

Deep neural network designs such as deep convolution neural networks (DCNN), recurrent neural networks (RNN), and long-short term memory (LSTM) are just a few of the deep neural network architectures that have been suggested and investigated for facial expression categorization (FEC). In order to identify and analyze the intricacies of human face expressions, those introduced in the literature have frequently needed great architectural complexity and computer capability. Handcrafted features such as Local Binary Patterns (LBP), dense optical flow, and Support Vector Machines (SVM) are examples of deeplearning alternatives[8].

The idea of facial expression recognition (FER) is crucial for a variety of uses, including picture retrieval, the neuroscience and psychology of human feeling, access control and surveillance, and the development of personality and young children. In addition to representing feelings, facial movements also reflect social contact, mental health, and physiological cues. According to psychophysiology research using facial electromyography (EMG), a method for facial emotions based on synchronized facial muscular activity in the viewer was demonstrated. Evolution in the a for mentioned disciplines is crucial for the study of and advancement of automatic face expression detection[3].

II. METHODOLOGY

In our paper, we will see that Emotion Net Nano, a extremely deep convolution Design ofa neural network for real-time facial emotion integrated scenario categorization. Emotion.Net Nano is built with a human collaboration method for machines in order to use both human and machine meticulousness.

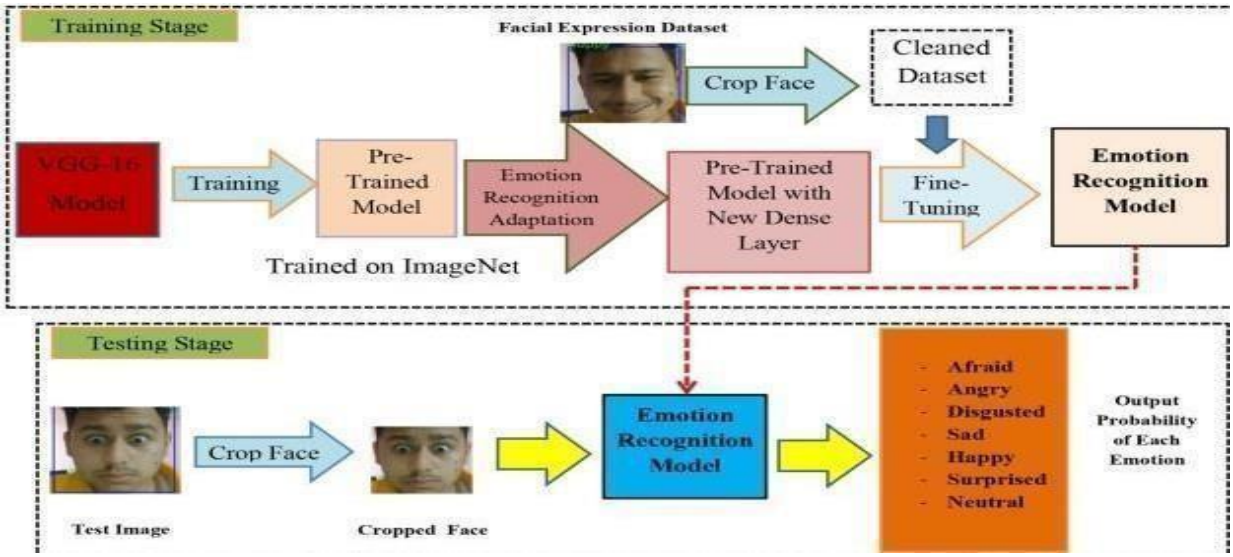


Fig 1: Showing the steps of facial expression recognition system during testing stage.

The machine – human collaborative design leveraged to create the proposed Emotion NetNano network design is comprised of two important designs.

1. Principal network Design
2. Machine-driven design exploration

Principled Network Design Prototyping

In the first design phase, an early network design prototype was built using human-centered design principles and served as a model for the next design phase of the machine-based survey. Previously shown to effectively perform various sensory tasks, the first prototype of the Emotion.Net Nano network design used architectural design elements in this study. Specifically, it has been shown thatthe inclusion of residual connections in a deep neural network design can successfully deal with both the vanishing gradient and the curse of dimensionality problems [8].

Furthermore, when contrasted to architectural or computational complexity, residual connections enable learning networks to be developed faster, simpler, and at a lower cost. As the network architecture's depth increases, each subsequent layer should outperform the previous layer. As a result, residual network architecture solutions that successfully address the FEC problem have been proven. In this study, the average

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

pooling function is used to achieve the final expression classification findings of the initial network design prototype, followed by a completely connected Soft Max activation layer.

Machine Driven Design

Taking after the introductory human-driven organize plan model stage, a machine-driven plan investigation stage was utilized to select the large scale design and smaller scale engineeringplans at the person module level to construct the ultimate Feeling Net Nano. Machine-driven planinvestigation was utilized to decide the ideal arrange engineering based on a set of imperatives set up by people utilizing generative union. The intriguing aspect of using an iterative approach to deal with the unconstrained optimization issue is that various generators are discovered along the way amid the optimization handle, with the capacity to create profound neural arrange structures with various trade-offs between engineering, computational, and exactness complexity.[8].

Background Study:

Human can recognize sentiments without any efforts, but machine recognition of facial expressions is extremely difficult. Some of the most common face expression tactics are:

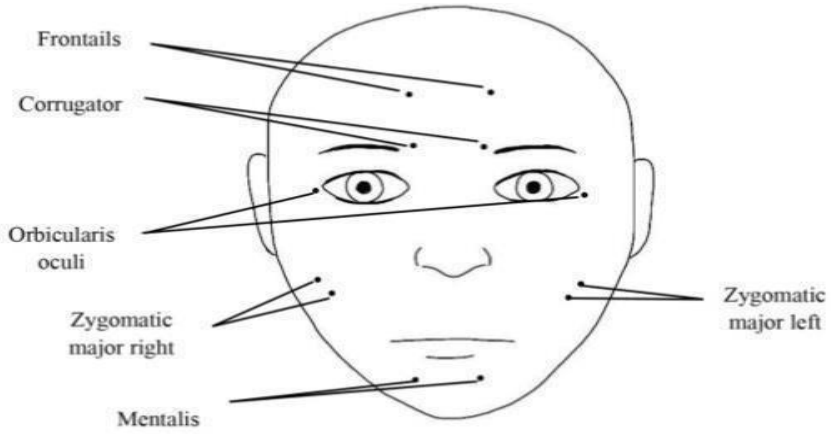


Fig 2: Illustration of main points of face and parts it's from where the detection will be done

Based on Statistical Motion:

This paper gives an invariant disturbance of noise and movement of head. Based on the zenicks moments of statistical movement.

Self – illumination dependent correction:

The facial recognition in this are described without marking them using the action unit to localize point . Face recognition is accomplished with use of the image skin and chromium. Mapping methods are used to map the mouth and the eyes to separate the face of the lack drop, hear – Cascaded differentiates skin and non-skin pixels[4].

Integrated Facial Expression: -

An instrument that can be identify face emotion is called the system of face detection. Facial detection recognition is used processing is used the processing useful information from an image can be extracted with of image processing.

The recognition system for facial expression consists of:

Data Acquisition: -

The elements that most people don't notice while they're looking for a new car or a house are the ones that stick out the most. The 2-D grey photo of face appearance scale identification is the most common. They would wish to be the same in future color images because of the low cost of color picture equipment. For picture acquisition, a cell phone or other digital device is employed.

Pre processing:-

Pre-processing is critical in the whole process. It's all about quality control and data minimization. It removes picture duplication without affecting the image's details. Pre- processing includes filtering and normalizing the picture, which results in consistent dimensions and a rotational image[10].

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

Segmentation: -

Segmentation splits the picture into relevant explanations. Picture segmentation is a method of separating a picture into homogenous and automated areas that fit various items in the image based on texture, edges, and intensity.

Extraction of Function: -

The removal of characteristics may be seen as part of "interest." They cover shape, movement, color, and facial image texture. The relevant image of knowledge is retrieved. When compared to the original picture extraction, the image detail, which gives an advantage in storage, is considerably diminished.

Categorization:

The extraction step is followed by the classification phase. The categorizing step defines, classifies, and aids in the differentiation of face pictures based on those classes. The classification step, also known as the feature selection stage, extracts and groups information based on certain parameters[10].

III. FACIAL EXPRESSION RECOGNITION (FER)

This technology includes system to recognize the fundamental human facial emotion detection with the use of on-going artificial intelligence algorithm, especially neural networks. The FER architecture comprises 3 steps: preprocessing, extraction of features and classification.

Pre-processing: To improve victory in profound learning tests centered on picture and flag preparing, information arrangement is as often as possible utilized. Confront arrangement, dark picture exchange, 2- D clamor evacuation versatile sifting, picture honing with sharp concealing, and information increment are all included in this area. Pre-processing improves the clarity of the incoming data (image) and reduces or eliminates duplication. The standard equation is then used to transform a picture input RGB of MN size into a grey image. Hear Cascade images library wasused to identify the circumference of the visage. These rectangular face emotions were then cut off and scaled. The pixel values of the photos were also converted into 64x64 gray scale images for use in neural networks. In a real-world scenario, picture data can be collected in a variety of situations, such asdistinct orientations, places, sizes, and accessibility. Consequently, with such unedited images, common pre-processing methods like standardization, chopping, and centralization enhance image recognition throughout any trial period[5].

Feature extraction: The initial step is to identify the face features in the image or video feed. Anger, disgust, fear, sorrow, pleasure, and amazement are the six primary emotions that the present study focuses on distinguishing as distinct traits. Utilizing the original data, the feature extraction approach generates new features. Reduced processing resource requirements without sacrificing important feature datasets are highly useful[5].

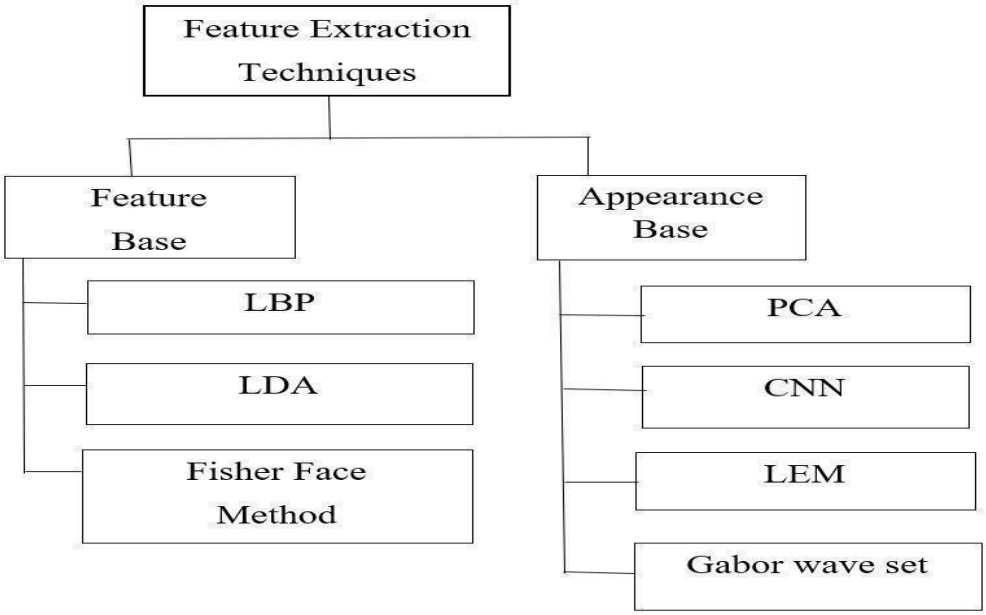


Fig 3: classification of feature extraction for facial expression recognition

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

Pre-processing In the feature extraction stage, data is modified to get the most precise representative characteristics. The retrieved data should offer insightful data that can be utilized to accurately predict the classification strategy. The outcome of the extraction procedure had an impact on the system's output. The two primary categories of conventional confront extraction methods are geometric methods, such as Dynamic Appearance Models (AAM), and appearance- based procedures, such as Gabor wavelet representation and Nearby Double Design (LBP). A geometric methodology considers different geometrical components, such as area, point, focuses of reference, and so on.

In an appearance-based strategy, the total input picture is inspected, the picture that most precisely speaks to the input picture is chosen, and highlights are recovered from it. serving as a profound extractor for extricating picture highlights. Taking after that, these characteristics are utilized to educated classical classification procedures like Back Vector Machine (SVM), Discriminant Investigation (LDA), and K- Closest Neighbors (K-NN). Several complex properties are created during the extraction process to explain the expression of face shape or texture changes.

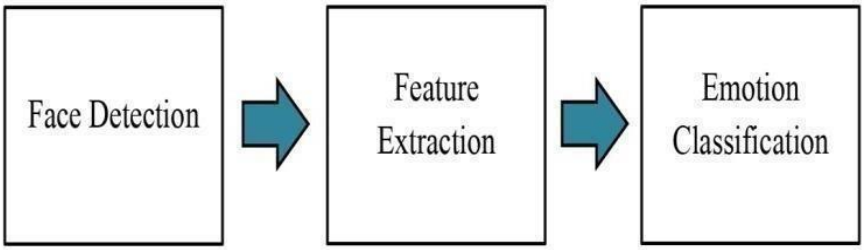


Fig 4: 3 Steps of emotion detection through face

Classification: is the concluding step of the FER technique, and it is here that the real there is mapping done for the action units' tagged emotions. Using the retrieved functions during the classification phase, a variety of strategies are employed to train the classifier. Recently, it has been demonstrated that deep learning (DL) techniques are an effective strategy for classifying images. Recurrent neural networks (RNNs), deep neural networks (DNNs), and CNNs are a few examples of deep neural network architectures that have been studied for their potential to recognize facial expressions [5].

This strategy employments an end-to-end approach to prepare a neural arrange structure with millions of parameters that learns redress highlights from enormous volumes of information without the utilization of hand- created highlights. In-depth characteristics or enactments can be utilized to input into standard classification calculations. In spite of the fact that any classification strategy can be utilized to achieve classification, the foremost utilized are SVM, K-NN, and LDA to extend acknowledgment precision. Taking after the include clarification, the highlights are connected to a classifier to recognize different facial expressions Consequently, the efficacy of the classifier is primarily decided by the consistency of the extracted characteristics. Each conventional classification has its own collection of features. SVM Kernels, for example, can distinguish distinct data sets by transforming several feature representations into higher dimensional space. SVM is also useful for object recognition and facial tracking. The LDA method, on the other hand, will decide the most effective way to distinguish between numerous groups [5,9].

Previous Related Works

Many studies and research have been done on the use of face emotion detection during the live classes be it online or offline. A study by Jian et al. 2020 showed that the technology used to detect facial emotion was able to detect emotions accurately that could be used to enhance the learning of students and increase productivity of teachers. And similarly many more studies were done before 2020 and after 2020 which also showed similar types results and few of them were also provided important feedback to faculties or teachers that if they change their way of teaching, add interactive way of teaching skills that helps in their job and increase the productivity that reflects on the learning of students and their results also **[**6].

Several approaches for FER have been investigated in recent decades. Traditional pioneer techniques first classify emotion based on feature values taken from the face picture. The most recent deep learning-based techniques, on the other hand, fulfill the FER job by merging the two processes into a single composite operating process. Many papers studied and contrasted current FER methodologies, and the most recent

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

research also incorporated deep learning- based techniques. The approaches employed in the well-known FER procedures are briefly outlined in the next sections..

Machine Learning-Based FER Approaches:

Programmed FER may be a intense issue within the field of counterfeit insights, especially within the machine learning subdomain. Different conventional machine learning approaches (such as K- closest neighbor and neural systems) are utilized to progress the FER issue. To begin with, the wavelet vitality include was included to the facial picture utilizing the ground- breaking FER method. Following that, highlights were recovered utilizing Fisher's straight separates (FLD), and feelings were categorized utilizing the KNN approach. In spite of the factthat PCA and NMF were utilized for include extraction, KNN was still utilized for classification in FER. Taking after the extraction of nearby binary pattern (LBP) histograms from different modest parcels of the picture, the feeling was recognized employing a direct programming (LP) approach [6].

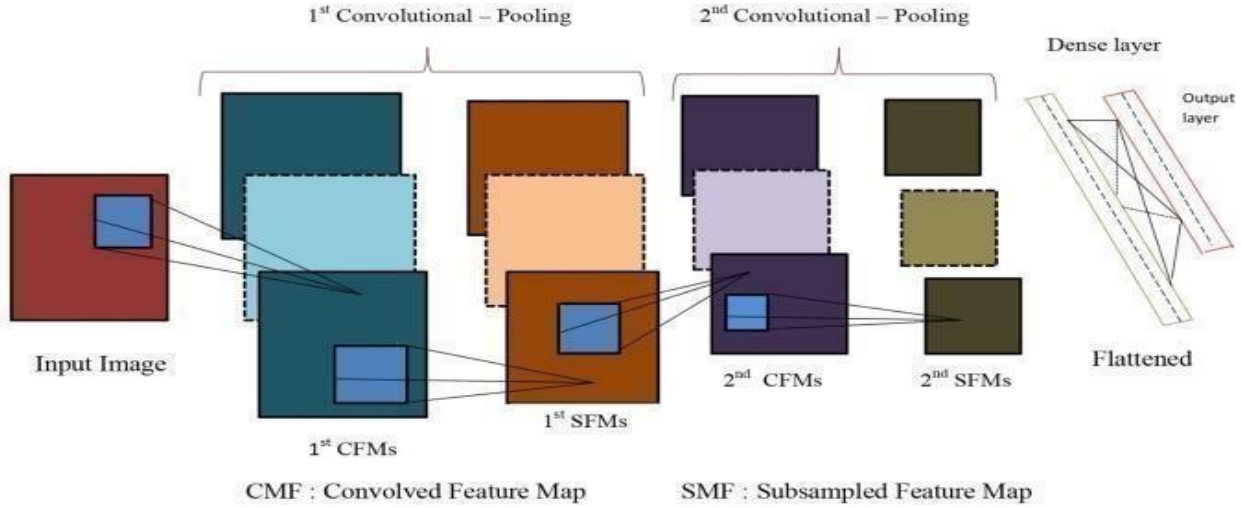
Deep Learning-Based FER Approaches:

Since CNN-based experiments were first reported in the literature, deep learning is a very recent machine learning method for FER. Zhao and Zhang used FER to merge a deep belief network (DBN) with a neural network (NN) to categorize emotion characteristics. The DBN is used for unsupervised feature learning. FER on independently collected face emotional pictures using a typical CNN architecture with two convolution-pooling layers. investigated a broader design that included two convolution-pooling layers and four inception layers. formed a group of 72 CNNs, each of which was trained using a different combination of the number of neurons in fully connected layers or a differentsize of filter in convolution layers [6].

Overview of CNN, Deep CNN Models and Transfer Learning (TL) &Convolution Neural Network (CNN):

This research is utilizing the TL method and pre-trained DCNN models. Several pre-trained DCNN models are evaluated to see which one is better for FER. To make the research self-contained, the subsections that follow provide an overview of CNN, potential DCNN models, and TL reasoning...

CNN is the most successful model for the image domain due to its inherent structure. An ACNN is made up of an input layer, multiple convolution-pooling hidden layers, and an output layer. Convolution, in its most basic form, is a mathematical operation performed on two functions to generate a third function that describes an altered shape of the original function.. A CNN's tiny (33, 55) kernel slides through the picture looking for meaningful patterns via convolution. Non-linear down sampling is a type of pooling. On the subsequent layer, a pooling layer aggregates non- overlapping regions from the preceding layer into a single value. Two convolution-pooling layers and a generic architecture for a typical CNN are shown in Figure 1.non-overlapping regions from the preceding layer into a single value. Two convolution-pooling layers and a generic architecturefor a typical CNN are shown in below figure.



The Generic architecture of a convolutional neural network with two convolutional- pooling layers

Fig 5: Explanation of CNN layers and its architecture

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

DCNN Models and TL Motivation

A DCNN requires high-dimensional pictures and contains numerous hiddenconvolution layers, making training and input highly difficult. Various DCNN models have various important connections and layouts in the convolution layers. Alex Net, a model that employs five-layer sin, was the first to achieve acceptable accuracy on Image Net. CNN is built on a comparable concept, but it used fewer parameters toreach an equivalent degree of accuracy. Smaller kernels have taken the place of larger ones. F Net obtained a comparable performance with only 1.3 million photos while Alex Net required 15 million images for training. VGG later offered a depth 16 model with 13 convolution layers and smaller kernels. VGG-19 is a model in this class that has 16 convolution layers.

In addition to a single skip link, Dense Net also contained dense skip connections across layers. This implies that every layer gets signals from layers below it and that every layer's output is utilised by levels below it. Input from one layer is combined with the channel concatenation of layers from layers that came before. Standard CNNs have L direct connections, whereas Dense Net has [L (L + 1)/2] direct connections. The network's information bottleneck is less severe since each layer has direct access to the layers above it. As a result, the model is exceptionally compute efficient and considerably thinner and smaller. The input to deeper layers will require a feature map since Dense Net blocks are created by combining feature maps[7,9].

Dataset

In way to measure the performance of the project based on facial detection in real time during class the basic emotions or expression like happy, sad, anger, surprise, disgust and neutral.

Each subject performed 150 instances for each expression in diverse scenarios. The age ranges of subject were 10 to 30 years and most them were males. About 350000 dataset are used in this project that are taken Kaggle. These were used at a speed of 15fps (frame per second

Fig 6: Examples of dataset emotion used in project

System Setup & Environment

The minimum hardware required for this type of project to run is a Intel core processor not less than i3 and a min 4gb ram, a proper internet connection, operating system window 8-10 orequivalent of 64 bit.

This type of project can be built with any programming language but here we developed on python on most popular Ides now days PyCharm because it is code completion, code navigation, refactoring, and debugging are just a few of the capabilities that this fully equippedide provides

Quantity Measurement

When we set the stratify parameter for training test split activity to ourlist of labels, the task will include part of each class, and ensure that this level is consistent with our training and validation data.

We have here classified the data according the some expression shown in below:

In this graph x line shows the expressions accuracy with respect to total number of datasets, Y line shows the number of datasets and bar chart shows the accuracy rate of expressions

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

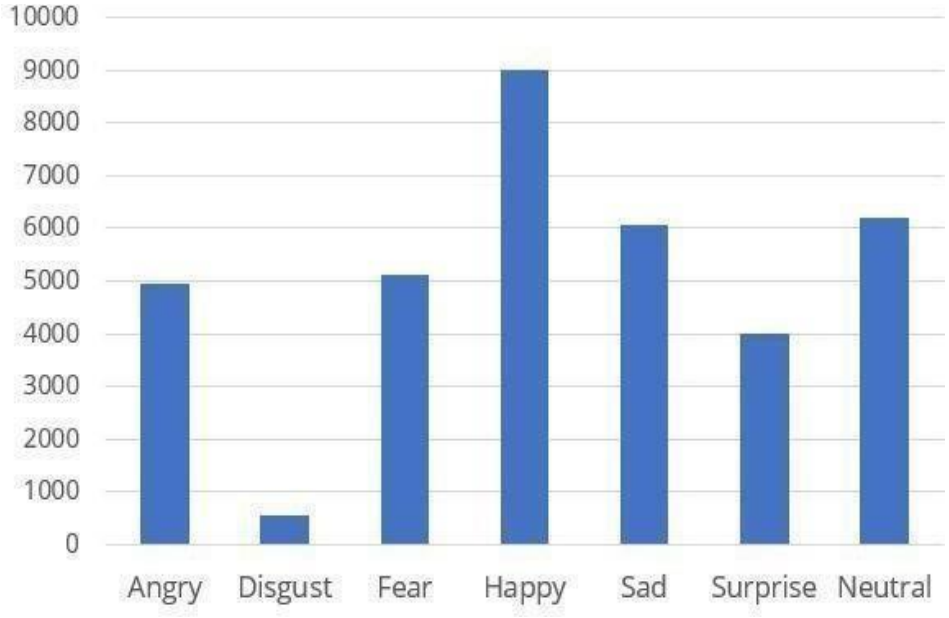


Fig 7: Quantity measurement data

Quality Validation

In the first image (surprise) it can detect the facial expression with help of eyebrows, forehead, mouth, chin and etc. it gives accuracy of 60% approx. of facial emotion detection.

In the second image (happy) it can detect the facial expression with help of eyebrows, forehead, mouth, chin and etc. it gives accuracy of 80% approx. of facial emotion detection.

In the third image (angry) it can detect the facial expression with help of eyebrows, forehead,mouth, chin and etc. it gives accuracy of 68% approx. of facial emotion detection.

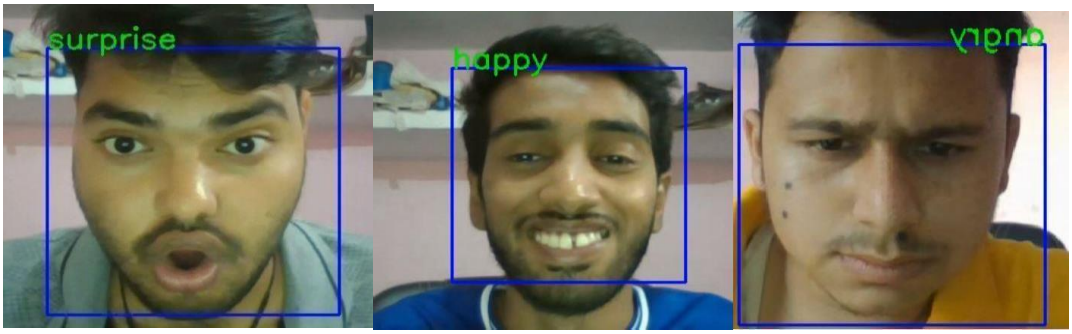


Fig 8 Quality measurement result

IV. RESULT

Here the result finally outputted work we got after the project successful completion oftraining stage. Here are some snapshot of the result of result that we tested on ourselves.

Fig 9: 1st testresults

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

Fig 10 .2nd Testresult

Future Trends

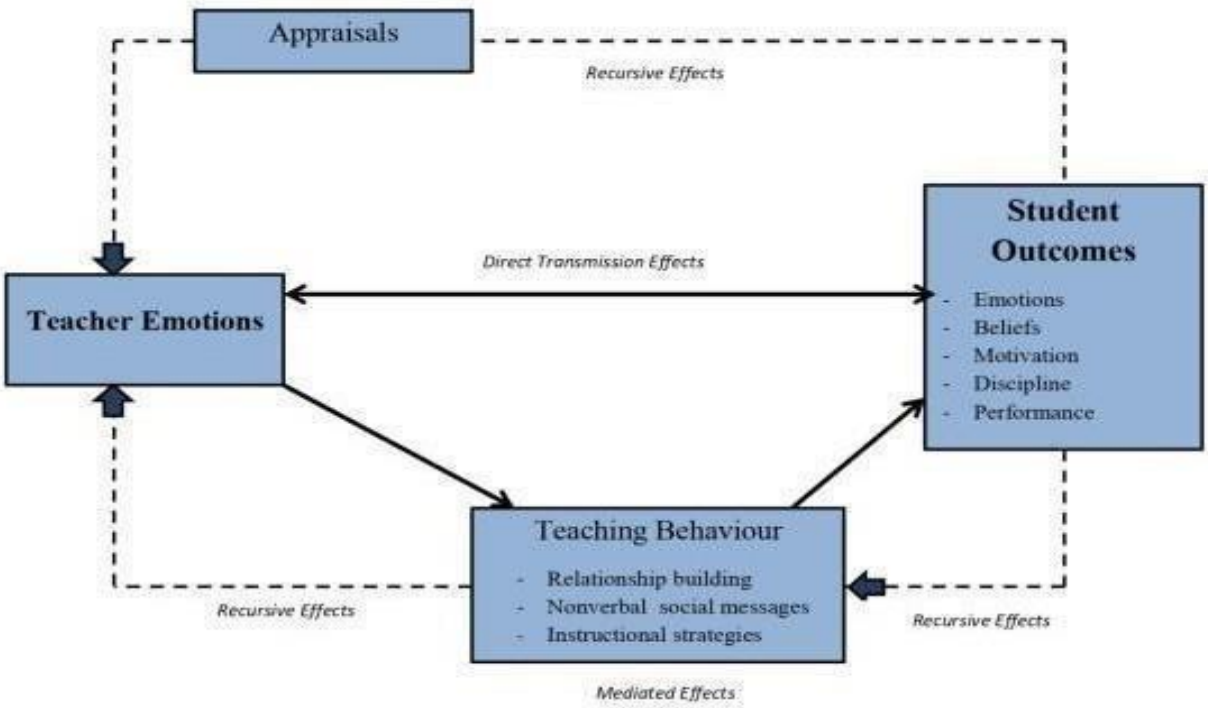


Fig 11: Effects of facial recognition in live class

V. CONCLUSION

Although emotion detection or recognition is one of the most effective ways to acquire emotional knowledge, it is typically limited to the six fundamental emotions plus neutral learning. Communication that includes emotional expressiveness improves the standard of human connection.

Furthermore, this research or article on facial recognition might contribute to greater feedback for society in the areas of education and communication between robot and AI interfaces in the future, which is already beginning to happen very quickly

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

VI. REFERENCES

[1] Illiana Azizan Faculty of Computer Science and Information Technology, University Putra Malaysia, Serdang,Selangor, Malaysia, August 2020

[2] James Ren Lee. etl Vision and Image Processing Lab, Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada, Darwin AI Corp., Waterloo, ON,Canada,13 January 2021, Doi: 10.3389/frai.2020.609673

[3] M. A. H. Ak hand. etl; Department of Computer Science and Engineering, Khulna University of Engineering & Technology, Khulna 9203, Bangladeshetl , 27April 2021

[4] Andrius Dzedzickis .etl Faculty of Mechanics, Vilnius Gediminas Technical University, J. Basanaviciaus g. 28, LT-03224 Vilnius,Lithuania; 21 January 2020

[5] Sharmeen M. Saleem Abdullah etl Duhok Polytechnic University, Duhok, Kurdistan Region,IRAQ, 15 April 2021,doi.org/10.30880/jscdm.2021.02.01.066

[6] Shuvendu Roy; S c h o o l of Computing, Engineering and Intelligent Systems, Ulster University, Londonderry

[7] Nazmul Siddique Graduate School of Science and Technology, Gunma University, Kiryu 376-8515, Japan

[8] Linda wang 1st ,2nd, Vision and Image Processing Lab, Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada, 2DarwinAI Corp., Waterloo, ON, Canada

[9] Byoung Chul Ko Department of computer Engineering, Keimyung university, Daegu 42601, Korea, 30 January 2018

[10] Chandra Bhushan Singh, Babu Sarkar, Pushpendra Yadav, Galgotias university, May 25, 2021

[11] Edeh Michael Onyema, Piyush Kumar Shukla, Surjeet Dalal, Mayuri Neeraj Mathur, Mohammed Zakariah, and Basant Tiwari Department of Mathematics and Computer Science, Coal City University, Enugu, Nigeria, 6 December 2021