JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Facial Expression Detection using Deep Learning

Hima Sainath Ballavarapu
Master of Technology (M.TECH)
Dept. of Computer Science and Systems Engineering
Specialization in Artificial Intelligence and Machine learning
Andhra University College of Engineering
Visakhapatnam, Andhra Pradesh, India

Dr Kompella Venkata Ramana (B.E, M.E, PHD)
Honorary Professor
Dept. of Computer Science and Systems Engineering
Andhra University College of Engineering
Visakhapatnam, Andhra Pradesh, India

Abstract:- Facial emotion recognition (FER) is an emerging and significant research area in the pattern recognition domain. In daily life, the role of non-verbal communication is significant, and in overall communication. Deep learning (DL) based emotion detection gives performance better than traditional methods with image processing. It discusses about the procedure of emotion detection, which includes basically three main steps: face detection, features extraction, and emotion classification. In this project a convolutional neural networks (CNN) based deep learning architecture and DL approaches VGG16 and ResNet50 for emotion detection from images. The performance of the proposed method is evaluated using three datasets Facial emotion recognition challenge (FERC-2013), MMA facial expression (MMAFEDB) and AffectNet. The highest accuracies achieved model is VGG16 having of 71.56 percentage for AffectNet dataset.

***Keywords*: Convolutional Neural Networks, Facial Expression, Deep Learning**

I. INTRODUCTION

Humans have always had the innate ability to recognize and distinguish between faces. Now computers are able to do the same. This opens up tons of applications. Face detection and Recognition can be used to improve access and security like the latest Apple IPhone does (see gif below), allow payments to be processed without physical cards — IPhone does this tool enable criminal identification and allow personalized healthcare and other services. Face detection and recognition is a heavily researched topic and there are tons of resources online. We have tried multiple open source projects to find the ones that are simplest to implement while being accurate. Face detection can be used for surveillance purposes by law enforcers as well as in crowd management. Facial expression

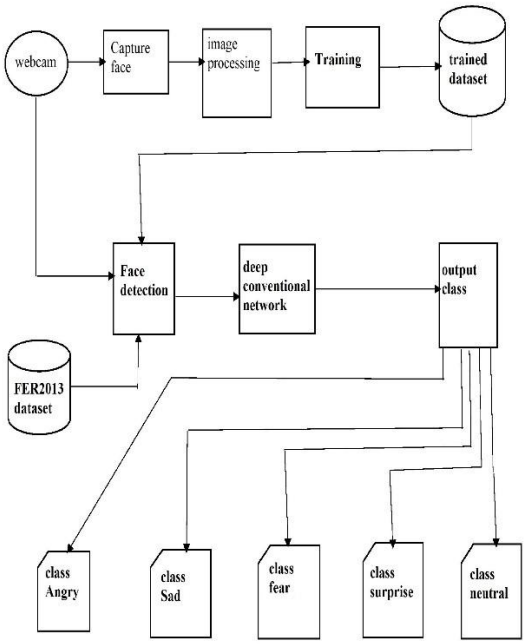
recognition is a biometric technique that expresses and analyzes human facial expressions, which enables computers to recognize and even understand human emotions. Expression recognition is originally used by psychologists to study people's psychology in crisis situations. With the integration with computer science, the technology has gradually made a series of attempts in the fields of human machine interaction in fatigue driving, smile detection, network video analysis, and mobile phone service. With the rapid development of service robots, it is very important to apply face emotional recognition to robot systems so that robots can understand human's emotion to better serve human beings. An automated expression recognition system receives an image or a video as input and output one of types of five emotions which are happy, anger, sad, surprise, and neutral.

II. LITERATURE REVIEW

Good fellow, I. J., Bengio, Y., & Courville, A. (2016). [1] Deep learning. MIT Press. This influential book provides a comprehensive introduction to deep learning, including Convolutional Neural Networks (CNNs). It covers the fundamental concepts, architectures, and training techniques that form the basis for modern facial expression detection systems. Liu, M., Li, X., & Huang, D. (2018). [2] Deep learning based facial expression recognition: A survey. Artificial Intelligence Review, 49(4), 1-16. This survey paper presents an overview of deep learning approaches for facial expression recognition. It reviews various CNN architectures, datasets, and performance evaluation metrics used in the field, highlighting recent advancements and challenges in the area of facial expression detection. Zhang, Z., Yan, C., & Liu, S. (2020). [3] A survey on deep learning for facial expression recognition. IEEE Transactions on Affective Computing, 11(4), 478- 500.This survey focuses on deep learning techniques specifically applied to facial expression recognition. It covers recent advancements in deep neural networks, including recurrent models and attention mechanisms, and discusses their impact on improving accuracy and robustness in facial expression detection. Lopes, A. T., de Oliveira, L. S., & da Silva, N. (2021) [4]. Deep learning-based facial expression recognition: A systematic review and comparative analysis. Neural Computing and Applications, 33(17), 11959-11980.This systematic review paper provides an in-depth analysis of deep learning-based facial expression recognition methods. It compares different CNN architectures, pre- processing techniques, data augmentation strategies, and training approaches, shedding light on the strengths and limitations of various methodologies Bartlett, M. S., Littlewort, G. C., Fasel, I., & Movellan, J. R. (2006). Fully automatic facial action recognition in spontaneous behavior. Journal of Multimedia, 1(6), 22-35. This paper presents a method for

automatic facial action recognition using a combination of Active Appearance Models (AAM) and Support Vector Machines (SVM). It demonstrates the potential for automated facial expression detection in real-world, spontaneous behavior scenarios. Shan, C., Gong, S., & McOwan, P. W. (2009) [5]. Facial expression recognition based on local binary patterns: A comprehensive study. Image and Vision Computing, 27(6), 803-816. Shan et al. conduct a comprehensive study on facial expression recognition using Local Binary Patterns (LBP). They explore various LBP-based feature extraction techniques and evaluate their performance on popular facial expression datasets, establishing LBP as a powerful method for facial expression detection. Zhang, J., Zheng, Y., & Sun, M. (2022)[6]. Facial expression recognition based on deep learning: A comprehensive survey. Journal of Ambient Intelligence and Humanized Computing, 13(9), 10937-10960. This comprehensive survey paper presents an overview of facial expression recognition techniques based on deep learning. It covers a wide range of topics, including facial feature extraction, dataset creation, network architectures, and performance evaluation, providing valuable insights for researchers and practitioners in the field. These selected papers provide a glimpse into the rich literature on facial expression detection. They cover seminal works, traditional methods, and recent advancements based on deep learning techniques, offering a solid foundation for further exploration and research in the area of facial expression detection and recognition.

III.SYSTEM DESIGN



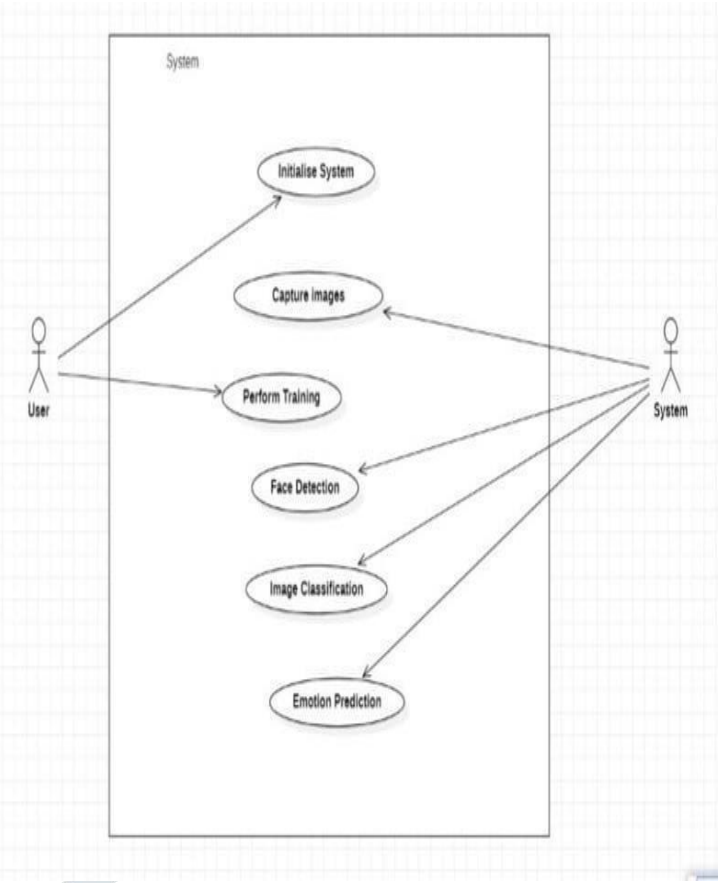
Proposed System Architecture

The proposed system architecture describes the workflow of the project we are working on. Ability to detect the location of face in any input image or frame. The output is the bounding box coordinates of the detected faces. It works by identifying and measuring facial features in an image. Facial recognition can identify human faces in images or videos, determine if the face in two images belongs to the same person, or search for a face among a large collection of existing images. By giving person's virtual face as an input through webcam it performs computations through trained dataset and it recognizes face once face got recognized it gives output based on the type of expression as output from the trained FERC 2013 trained dataset.

IV.UML DIAGRAMS

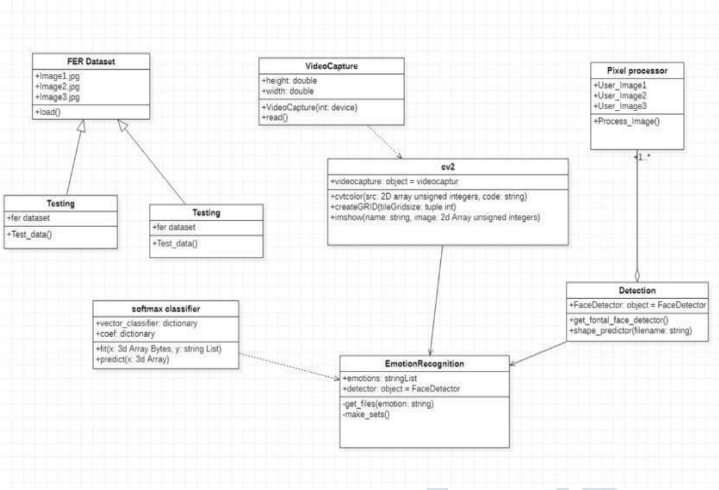
Use case Diagram:

A Use Case consists of use cases, persons, or various things that are invoking the features called as actors and the elements that are responsible for implementing the use cases. Use case diagrams capture the dynamic behavior of a live system. It models how an external entity interacts with the system to make it work. Use case diagrams are responsible for visualizing the external things that interact with the part of the system. Use cases are used to represent high- level functionalities and how the user will handle the system.



Class diagram:

Class diagram is a static diagram. It represents the static view of an application. Class diagram is not only used for visualizing, describing, and documenting different aspects of a system but also for constructing executable code of the software application. Class diagram describes the attributes and operations of a class and also the constraints imposed on the system. The class diagrams are widely used in the modelling of object-oriented systems because they are the only UML diagrams, which can be mapped directly with object-oriented languages. Class diagram shows a collection of classes, interfaces, associations, collaborations, and constraints. It is also known as a structural diagram.



IV.PROPOSED METHODOLOGY

Proposed Methodology

The proposed system architecture describes the workflow of the project we are working on. Compare multiple faces together to identify which faces belong to the same person. This is done by comparing face embedding vectors. A face analyzer is software that identifies or confirms a person's identity using their face. It works by identifying and measuring facial features in an image. Facial recognition can identify human faces in images or videos, determine if the face in two images belongs to the same person, or search for a face among a large collection of existing images. Biometric security systems use facial recognition to uniquely identify individuals during user on boarding or logins as well as strengthen user authentication activity. Mobile and personal devices also commonly use face analyzer technology for device security.

Emotion Detection

Classifying the emotion on the face as happy, angry, sad, surprise and neutral.

CNN Architecture

The networks are program on top of keras, operating on Python, using the keras learn library. This environment reduces the code's complexity, since only the neuron layers need to be formed, rather than any neuron. The software also provides real-time feedback on training progress and performance, and makes the model after training easy to save and reuse. In CNN architecture initially we have to extract input image of 48*48*1 from dataset FERC-2013 and MMAFEDB. The network begins with an input layer of 48 by 48 which matches the input data size parallelly processed through two similar models that is functionality in deep learning, and then concatenated for better accuracy and getting features of images perfectly. This models contains convolutional layer with 64 filters each with size of [3*3], followed by a local contrast normalization layer, maxpooling layer, followed by one more convolutional layer, max pooling, flatten respectively. After that we concatenate two similar models and linked to a softmax output layer which can classify seven emotions. We use dropout of 0.3 for reducing over-fitting. It has been applied to the fully connected layer and all layers contain units of rectified linear units (ReLu) activation function. Output Layer units of 5.

VGG16 Architecture

VGG16 is a convolution neural net (CNN) architecture which was used to win ILSVR(Imagenet) competition in 2014. It is considered to be one of the excellent vision model architecture till date. Most unique thing about VGG16 is that instead of having a large number of hyper-parameter they focused on having convolution layers of 3x3 filter with a stride 1 and always used same padding and maxpool layer of 2x2 filter of stride 2. It follows this arrangement of convolution and max pool layers consistently throughout the whole architecture. In the end it has 2 FC(fully connected layers) followed by a softmax for output. The 16 in VGG16 refers to it has 16 layers that have weights. This network is a pretty large network and it has about 138 million (approx.) parameters

ResNet50 Architecture

ResNet, short for Residual Networks is a classic neural network used as a backbone for many computer vision tasks. This model was the winner of ImageNet challenge in 2015. The fundamental breakthrough with ResNet was it allowed us to train extremely deep neural networks with 150+layers successfully. Prior to ResNet training very deep neural networks was difficult due to the problem of vanishing gradients.

Datasets

- FERC-2013
- MMAFEDB
- Affect Net

V.EXPERIMENTAL RESULTS

Tabulated results

Qualitative assessment of models for emotion detection

The above tabulated results are obtained by training with different models and different algorithms.Among these VGG16 – AffectNet dataset has been acquired with highest accuracy of 71%. The accuracy of the model depends on the picture quality and resolution of the data sets

Model and Dataset	Validation accuracy (%)	Validation Loss	Computation time per step (m sec)
CNN - MMAFEDB	**56**	2.03	2 m sec
VGG16 - MMAFEDB	**41**	1.51	182 msec
ResNet50 - MMAFEDB	**37**	1.60	268 msec
CNN – FERC2013	**63**	**1.35**	5 msec
VGG16 - AffectNet	**71**	**1.03**	2 msec

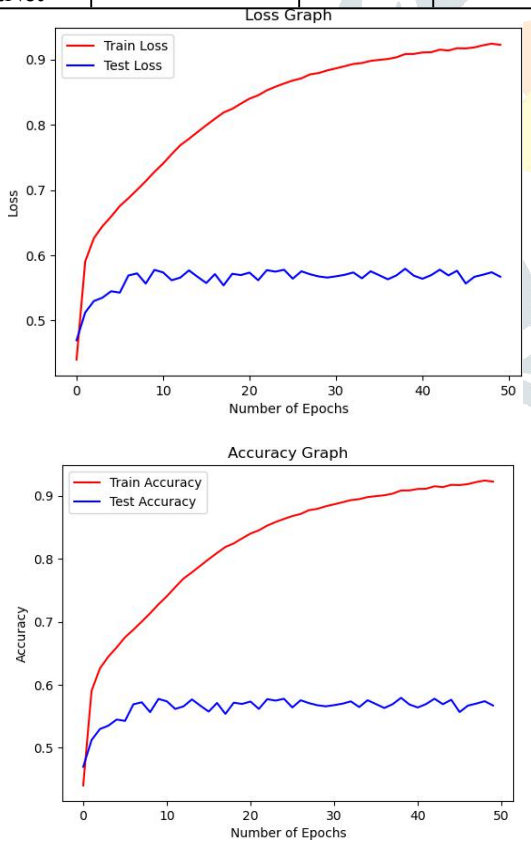


Figure: Train , Test Accuracy and loss of CNN model for MMAFEDB dataset

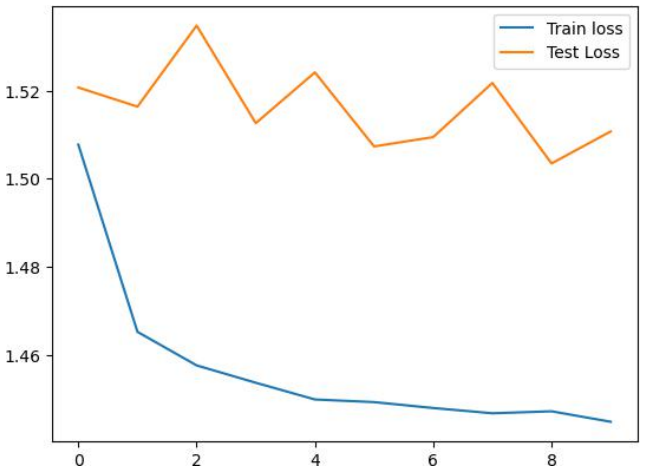


Figure: Train , Test Accuracy and loss of VGG16 model for MMAFEDB dataset

The above two graphs descirbes about the train and test accuracy of the CNN model for MMAFEDB dataset and VGG16 model for MMAFEDB dataset. The following figures which comprises of graphs shows how there is a variation across the accuracy and loss figures across the two models and it visualises about how the given algorithm woks on .

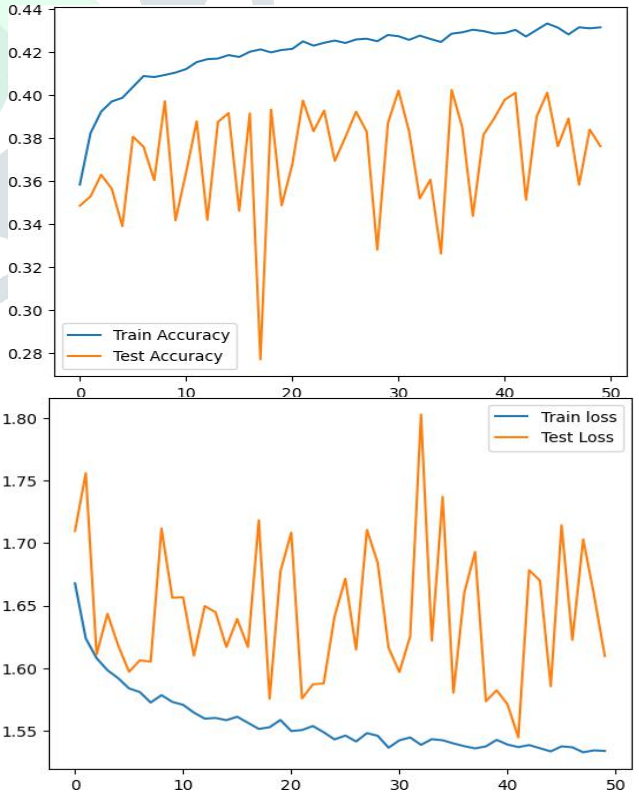


Figure: Train , Test Accuracy and loss of ResNet50 model for MMAFEDB dataset

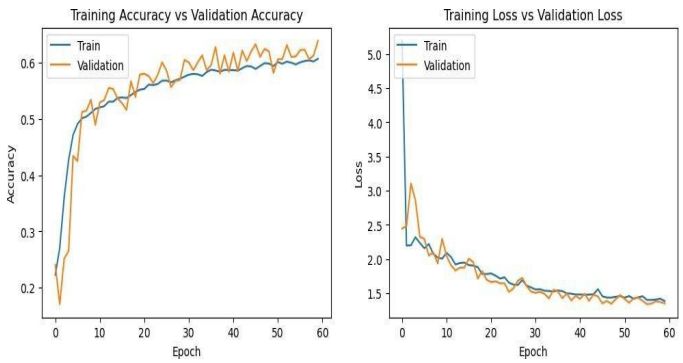


Figure: Train , Test Accuracy and loss of CNN model for FERC2013 dataset

The above graphs shows the train and test accuracy and loss of ResNet50 model for MMAFEDB dataset similarly in the second graph the train and test accuracy of CNN model for FERC2013 dataset is displayed.

Similarly in the below graphs it represents about the train,test accuracy and loss of AffectNet dataset.

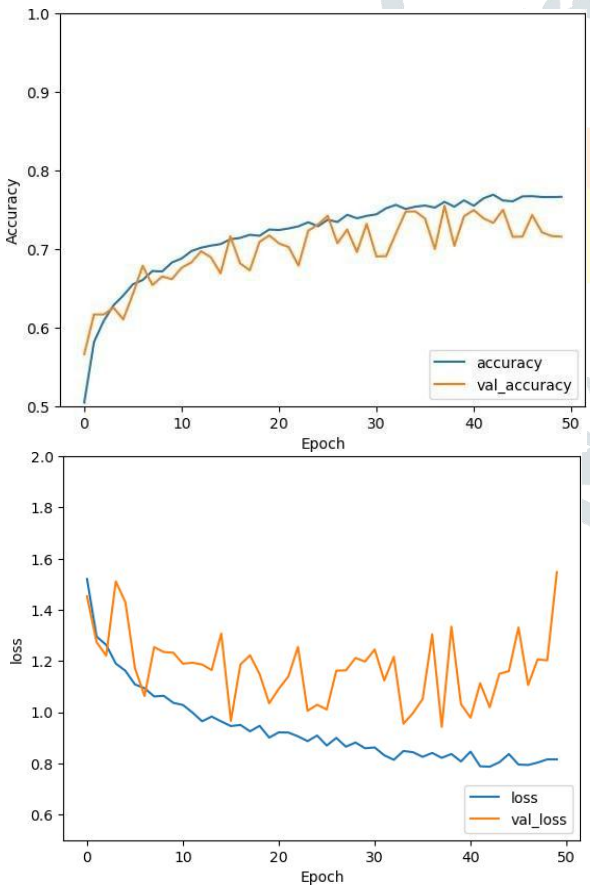


Figure: Train ,Test Accuracy and loss of VGG16 model for AffectNet dataset

VI.CONCLUSION

In this Project, we have proposed a deep learning based facial emotion detection method from image. Using different DL models using three different datasets, MMAFEDB, AffectNet and FERC-2013. The performance evaluation of the proposed facial emotion detection model is carried out in terms of validation accuracy, computational complexity, detection rate, learning rate, validation loss, and computational time per step. Analyzed models using trained and test sample images, and evaluate their performance compare to previous existing model. Results of the experiment show that the model proposed is better in terms of the results of emotion detection. The highest accuracies achieved model is VGG16 having of 71.56 percentage for AffectNet dataset. We can use this models to implement an applications, in this project we taking only 5 facial expressions (Anger, Surprise, Sad, Happy and Neutral). For future proposed we can add more facial expressions. The facial expression recognition system presented in this research work contributes a resilient face recognition model based on the mapping of behavioral characteristics with the physiological biometric characteristics. The physiological characteristics of the human face with relevance to various expressions such as happiness, sadness, fear, anger, surprise and disgust are associated with geometrical structures which restored as base matching template for the recognition system.

The behavioral aspect of this system relates the attitude behind different expressions as property base. The property bases are alienated as exposed and hidden category in genetic algorithmic genes. The gene training set evaluates the expressional uniqueness of individual faces and provide a resilient expressional recognition model in the field of biometric security.

The design of a novel asymmetric cryptosystem based on biometrics having features like hierarchical group security eliminates the use of passwords and smart cards as opposed to earlier crypto systems. It requires a special hardware support like all other biometrics system. This research work promises a new direction of research in the field of asymmetric biometric cryptosystems which is highly desirable in order to get rid of passwords and smart cards completely.

Experimental analysis and study show that the hierarchical security structures are effective in geometric shape identification for physiological traits.

VII.Future Enhancements:

It is important to note that there is no specific formula to build a neural network that would guarantee to work well. Different problems would require different network architecture and a lot of trail and errors to produce desirable validation accuracy. **This is the reason why neural nets are often perceived as "black box algorithms."**

In this project we got an accuracy of almost 70% which is not bad at all comparing all the previous models. But we need to improve in specific areas like-

- **Number and configuration of convolutional layers**
- **number and configuration of dense layers**
- **Dropout percentage in dense layers**

But due to lack of highly configured system we could not go deeper into dense neural network as the system gets very slow and we will try to improve in these areas in future.

We would also like to train more databases into the system to make the model more and more accurate but again resources becomes a hindrance in the path and we also need to improve in several areas in future to resolve the errors and improve the accuracy.

Having examined techniques to cope with expression variation, in future it may be investigated in more depth about the face classification problem and optimal fusion of color and depth information. Further study can be laid down in the direction of allele of gene matching to the geometric factors of the facial expressions. The genetic property evolution framework for facial expressional system can be studied to suit the requirement of different security models such as criminal detection, governmental confidential security breaches etc. .

REFERENCES

[1].Mehmet Akif OZDEMIR1 , Berkay ELAGOZ2 , Aysegul ALAYBEYOGLU SOY3 , andAydin AKAN4, "*Deep Learning Based Facial Emotion Recognition System*", 978-1-7281- 8073-1/20 – IEEE 2020.

[2].Paras Jain, M Murali, Amaan Ali, "*Face Emotion Detection Using Deep Learning*", 978- 1-6654-2642-8 – IEEE 2021

[3]MMA FACIALEXPRESSION| Kaggle

[4]FER-2013 | Kaggle
www.kaggle.com/datasets/msambare/fer2013

[5]AffectNet-Sample | Kaggle
www.kaggle.com/datasets/mouadriali/affectnetsample

[6]. https://towardsdatascience.com/face-detection-recognition-and-emotion- detection-in-8- lines-of-code-b2ce32d4d5de

[7].A literature survey on Facial Expression Recognition using Global Features by Vaibhavkumar J. Mistry and Mahesh M. Goyani,International Journal of Engineering and Advanced Technology (IJEAT),April,2013[http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.6 45.5162&rep=rep1&typ e=pdf].

[8]. **Convolutional Neural Networks (CNN) With TensorFlow** by Sourav from Edureka[https://www.youtube.com/watch?v=umGJ30-15_A].[9]. **A Survey on Facial Expression RecognitionTechniques by Swati** Mishra(Chhattisgarh Swami Vivekanand Technical University, Department of Computer Science& Engineering, Raipur Instituteof Technology, Mandir Hasaud, Raipur, Chhattisgarh,India) and Avinash Dhole(Assistant Professor, Department of ComputerScience & Engineering, RaipurInstitute of Technology, Chattisgarh Swami Vivekanand and Technical University,MandirHasaud, Raipur,Chhattisgarh,India), International Journal of Science and Research (IJSR),2013