EasyChair Preprint

№ 13204

Face Sentiment Prediction Using Deep Learning

S.N Chandra Shekhar, Abhijith Bandari,
Sai Prasanna Perumandla and Sunia Sultana

May 6, 2024

FACE SENTIMENTAL PREDICTION USING DEEP LEARNING

Dr.S.N.Chandra Shekhar
Assistant Professor, ECE Department
snchandrashekhar@gmail.com SNIST, HYDERABAD

B . Abhijith
ECE(SNIST)
20311a04x3@sreenidhi.edu.in

P . Sai Prasanna
ECE(SNIST)
20311a04z6@sreenidhi.edu.in

Sunia Sultana
ECE(SNIST)
20311a04ak@sreenidhi.edu.in

Abstract: ***In this paper we are going to discuss about real-time image predicting human facial emotions. This can done by using few software techniques. Here we are going to use Deep Learning Techniques to detect Human Face and Emotion. This will be very helpful for Human-Computer Interaction which leads to improved technology in many Technical ways. By using GUI i.e. Graphical User Interface, we can build a face sentimental recognition Technique. The main goal of this project is to bring to employ deep learning and machine learning techniques to create GUI for face sentimental recognition. This will be very helpful in medical , marketing , psychology , etc. fields. By improving machine learning techniques for face detection and emotion classification , this application allows user to upload images containing humans through which this application going to detect face and classify emotion included in that face. Here , we use Haar Cascade to detect face and Xception algorithm to detect emotion. Firstly we are starting the process by training the algorithm and later by feedback mechanism it will be predicting the next output based on previous output. In this way prediction of emotion does occurs. Haar Cascade will be helpful in detecting human face and Xception will useful to train and test the data i.e. to recognize the facial emotion by the image provided. The main aspect of this project to detect emotion in human by real-world image which helps machines to know the emotion of human using Xception.***

Keywords: Face-Emotion Detection, Human-Computer Interaction, Deep Learning, Machine Learning, Xception, Haar Cascade.

I. INTRODUCTION

Nowadays Face emotion plays a very crucial role in many aspects of day-to-day life. So, detection of emotion had became very important in our daily life. In many aspects like health care, marketing , business , psychology etc.

We humans can easily identify the emotion of a human just by looking at their face. But, machines cannot perform everything as we humans did. So, here with the help of softwares we are going to make the machines recognize human expressions by real world images of human faces. In the world of Digital interactions and Technology enhancement, we must build the interaction between the user and machine which is user-machine interaction , here such software will be very helpful for us for machine learning. Machine Learning is one such study in which artificial intelligence was included , which was aimed to build a strong bridge between human and machine. This study mainly concerns about human computer interaction and to bring human intelligence to computers which will be very helpful for us to make our activities easier and faster.

One such aspects is this face sentimental analysis using machine learning. Here we are going to use Haar Cascade and Xception which is a deep learning algorithm with some machine learning techniques like k-nn and svm. These machine learning techniques will be helpful in emotion classification whereas xception will be useful for training and testing the data.

In the recent years , we can observe many experiments were held on this face sentimental analysis to improve face sentimental prediction and to learn more about face recognition. Many techniques were used to get the face recognition by computer. They combined many methods and techniques to get the most accurate results of the facial recognition.

Comparing with the early days of machine learning , now a days we can observe a very great difference of machine learning. Here we can observe some very new techniques created and utilized inorder to make our work easy. Even we have advanced technology to carry on the experiments that we need which takes some hours in past , now they can be completed only in very much less time. All this techniques are developed from the zeal of learning. The zeal to learn more and to implement more brings out such wonderful techniques.

So, In order to know more about this machine learning techniques and algorithms people keep on experimenting on it. And made many new techniques which gives us more accuracy values and some techniques that takes much less time and some takes less time and gives more accuracy.

This motivates us more to learn this deep learning techniques along with machine learning techniques to detect the emotion on the human provided.

II . LITERATURE SURVEY

Facial recognition techniques served as the foundation for Donato's approach (Yang & Kriegman, 2002). The facial image includes identity information in addition to the expression. Retrieving the identifying information from the image is useful to perform a comprehensive study of the expression. Images from the movie were combined to generate a new image, and optical flow estimates and other spatial studies were utilized to classify the expressions. Again, Pantic and Rothkrantz's technique required matching the image to point-based face models to identify the facial movement units (Action Units), which were subsequently classified into various emotional states (Pantic & Rothkrantz, 2001).In his study, Colmenarez also models faces using feature points (Colmenarez et al., 1999). It should be highlighted, nonetheless, that one disadvantage of these approaches is their excessive dependence on the accuracy of the feature points. Numerous studies have focused on the tracking and classification of facial expressions. Dhavaliker and Kulkarni (2014) state that face detection, facial feature extraction, and facial expression recognition are the three basic procedures usually involved in these investigations. When determining the facial region, characteristics such as skin tone and the geometric shape of the face are taken into consideration

III . EXISTING SYSTEM

SVM, a supervised machine learning algorithm, is used to solve problems with regression and classification. The SVM method is a binary classifier used in classification applications. The SVM algorithm represents each item of data as a point in n-dimensional space with a given coordinate value. It then ascertains which hyperplane splits the two classes the most effectively. The method classifies the data according to which side of the hyperplane it is on (Khoong, 2021). k-NN is one algorithm for supervised machine learning. Thus, the program uses the labeled dataset. The class of the new observation is ascertained using the classes of the k closest observations. The class that gets the most votes from its k closest neighbors determines how the new observation is categorized (Srivastava, 2018). CNN is a deep-learning network that processes images very well. By comparison, it emerges from the visual cortex. The CNN architecture processes the image across multiple layers. The layers that are commonly employed are convolution, pooling, flattening, completely connected, and softmax layers. Several architectures can be produced by arranging these layers in different ways (Verma &Verma, 2020). Çınari (2018) states that Alexnet is the first large-scale CNN architecture to win the Imagenet competition.

IV .DISADVANTAGES OF EXISTING SYSTEM

To properly learn hierarchical features from images, CNNs need a lot of labeled training data, which can be difficult or expensive to get in some fields.

Training deep CNNs can be computationally intensive and time-consuming, especially on hardware with limited computational resources.

CNNs may suffer from overfitting, especially when training on small datasets or when the model architecture is too complex relative to the available data.

V .PROPOSED SYSTEM

The software used in this study was Keras-TensorFlow, one of the deep learning libraries available for the Python programming language. Figure 1 displays the proposed method's block diagram. When the software is launched, it takes images from the laptop and performs preprocessing before detecting faces. The HaarCascade is used to detect Human faces and The Xception method is then used to extract features. Convolutional neural networks that are specifically designed are used to automatically perform the features. Seven universal emotions happiness, sadness, anger, disgust, surprise, neutral, and fear can be categorized by taking facial features.

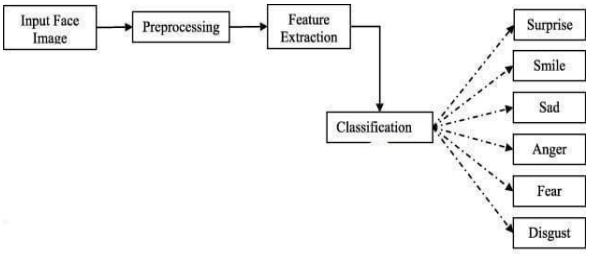


Figure 1. The suggested facial expression-based image processing system for emotion recognition

VI . FUNCTIONAL REQUIREMENTS

Functional requirements for a face sentiment detection systemincludes the actions we take in the project and methods we use. Here are the functional requirements for such a system:

Image Upload:

- Here Users should upload their image that contains a human face.

Face Detection:

- Here we are going to get the accurate face emotion detection of the image that been uploaded.

Emotion Recognition:

- The main aim of this project is to detect emotion of the human face recognized in the image. Emotions include sadness, happy, fear , surprise , disgust and netural.

Real-Time Processing:

- The Emotion detection will be occur by real-time images or photographs that been uploaded.

Feedback Display:

- The output of the present input will be based on the experience earned by

previous output.

Song Recommendation:

- After detecting the emotion, based on the emotion detected a song will be recommended.

VII . DATA COLLECTION

Data collection for a face sentiment detection system involves assembling a diverse and comprehensive dataset of images depicting facial expressions representing various emotions. To begin, this procedure specifies the precise emotional categories—including joy, sorrow, fury, surprise, fear, disgust, and neutrality—that the system is to identify. Images are sourced from multiple channels, including online databases, social media platforms, public datasets, or through in-house data collection efforts. It's essential to ensure the dataset encompasses a wide range of individuals, spanning diverse demographics, ages, genders, ethnicities, and cultural backgrounds. Additionally, images captured under different lighting conditions, angles, and environmental settings contribute to the dataset's variability and robustness. Once collected, each image undergoes annotation, where it is labeled with the corresponding emotion expressed by the individual in the image. Annotation can be performed manually by human annotators or through automated tools, with a focus on maintaining consistency and accuracy in labeling. The dataset may be augmented using methods like as rotation, translation, scaling, flipping, and noise addition to promote variety and inhibit overfitting.

VIII. HARDWARE REQUIREMENTS:

- **Processor :** Pentium –IV
- **RAM** : 4 GB (minimum)
- **Hard Drive: 20 GB**
- **Key Board :** Standard Windows Keyboard
- **Mouse** : Two or Three Button Mouse
- **Monitor** : SVG

IX . SOFTWARE DESIGN

The goal of this project is to create a real-time facial recognition system using the Python programming language. The software was installed on a laptop computer with strong hardware because it was compatible with all three databases. The application, various settings, and tests were run multiple times on an Acer laptop with an 8-core Intel Core i5 processor running Python in the PyCharm to get the best results. Python is an object-oriented, high-level, interpretive, modular, interactive programming language. The language's simple grammar, which is built on indentations, makes it simple to learn and remember. This makes it a language that allows one to begin programming without wasting time on syntax nuances (Python Software Foundation, 2012). Except for the face recognition stage, all other phases utilize only CPU cores.

Keras is a high-level neural network library that runs on top of TensorFlow, while OpenCV is a computer vision library. TensorFlow is an open-source end-to-end platform.

Software library for several tasks related to machine learning (Terra, 2022). OpenCV is a library designed specifically for computer vision techniques. TensorFlow is a framework for machine learning problems. TensorFlow can also be used for more general problems like regression, classification, and clustering, but it is particularly useful for image recognition. Written in Python, A high-level neural network API that works with Tensorflow is called Keras. It states that Tensorflow, a machine learning framework, has been widely accepted by the research and industrial community for deep learning tasks. (Mokhtari, 2021) Tensorflow is used to write the majority of deep learning models that are publicly available. Simple face detection methods such as SVM and K-NN can be tested with OpenCV.

X . USE CASE OUTLINE

According to UML(Unified Modelling Language) , use case diagram is a diagram that gives the overall outline of the project and provides knowledge about the steps that been utilized.

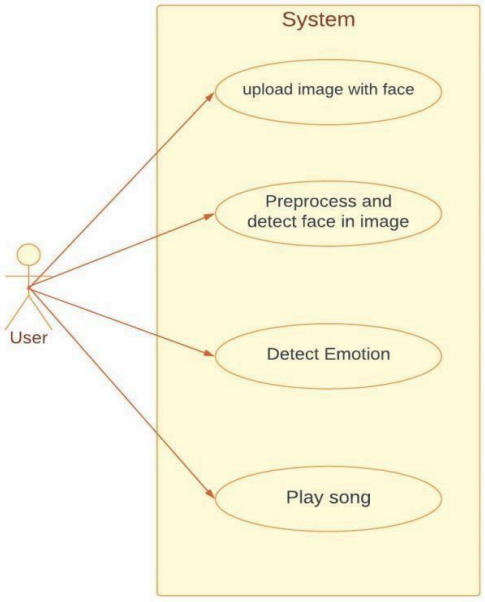


Figure 2: Predicts the use case diagram of project

XI . ACTIVITY GUIDE

Activity diagram depicts the activities that been occurred in the project. It is one of UML diagrams , which provides information about the activities that been going that specific project. Step-by-step ,it will depict entire activity of the project.

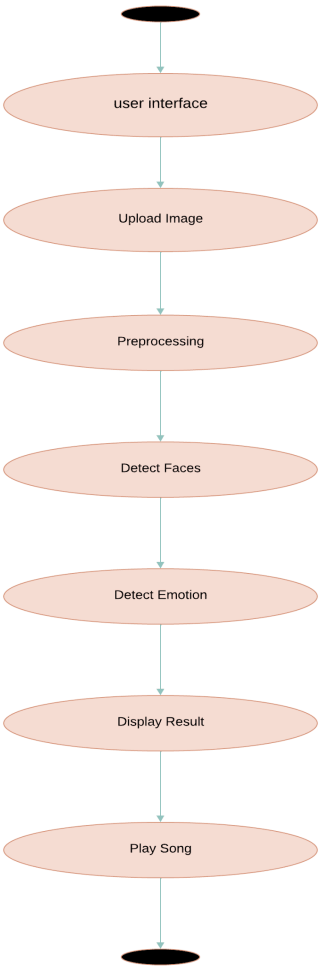


Figure 3: The Activity Diagram of the following project is predicted.

XII . SYSTEM IMPLEMENTATION

HAAR CASCADE CLASSIFIER

Organize the 6000 features into distinct classifier stages and apply each stage individually rather than applying them all at once on a window. (First few phases typically include extremely few features). Throw away a window if it doesn't pass the first test. We disregard any further features that may be present. Apply the second set of features and carry on with the procedure if it passes. A face area is the portal through which all phases pass." Face Recognition via Haar Cascades.

Xception Algorithm:

Xception is an extremely effective deep-learning network for image processing. It develops from the visual cortex through analogy. The image is processed by going through several layers in the Xception architecture. Convolution, pooling, flattened, fully connected, and softmax layers are the layers that are frequently used. By stacking these layers, various architectures can be created (Verma &Verma, 2020). According to Çınar (2018), Alexnet is the first large-scale Xception architecture to win the Imagenet competition. Three fully connected layers and five convolution layers make up the Alexnet architecture. Approximately 1.3 million photos with 1000 classes were used to train it. A different Xception architecture is the Residual Network (ResNet). The main essential element of the Resnet design is residual blocks. The remaining links are employed in this architecture to avoid the vanishing gradient issue. During neural network training, different convolution and pooling processes result in losses in the feature map. A thick net architecture has been proposed as a way to prevent this. I Xception's architecture is composed of dense bricks. The layers of the blocks are connected in a complex way. Lastly, InceptionV3's Xception architecture was used. This design, according to Islam and Al-Murad (2017), provides a first model that combines multiple convolutional filters of different sizes to generate a new filter. This model requires fewer parameters to be trained. This reduces the computational complexity of the model as well.

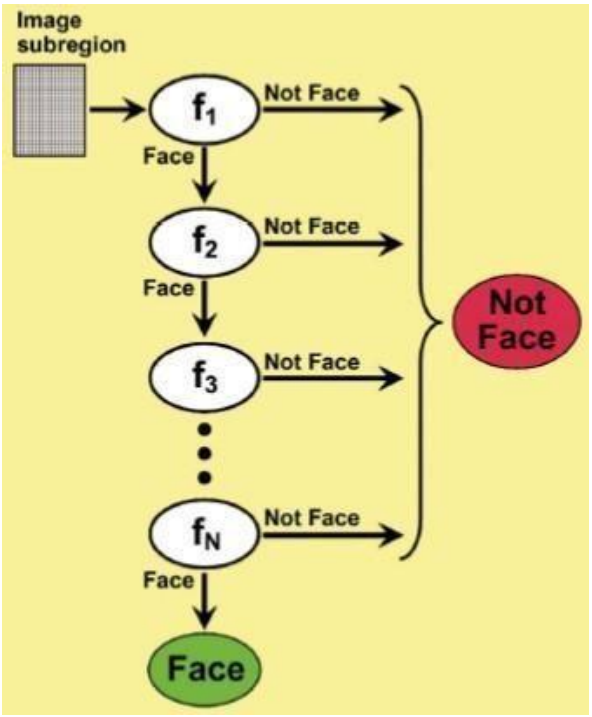


Figure 4 This Figure illustrates the working of the Haar Cascade classifier.

XIII . RESULTS AND DISCUSSION

The overall result of this project getting the emotion of the image that uploaded which consists human face in it. Here, before predicting the emotion , it also predicts the number of faces included. After emotion detection project will recommend a song according to the emotion predicted. Using the Fer2013 dataset further emphasizes how crucial ethical issues are when handling data and developing models. Safeguarding the confidentiality and integrity of people's facial data in the dataset is crucial, necessitating strong security and anonymization protocols. Overall, the integration of the Fer2013 dataset into the face sentiment detection project yields promising results, paving the way for further advancements in emotion recognition and machine learning. Through the use of this dataset, the research advances knowledge on face emotion detection and its possible uses in a variety of industries, such as social robots, healthcare, and

human-computer interaction. The confusion matrix from the study is also shown in Figure.

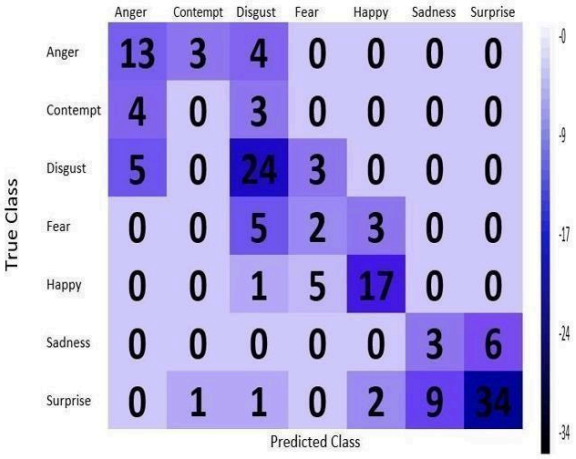
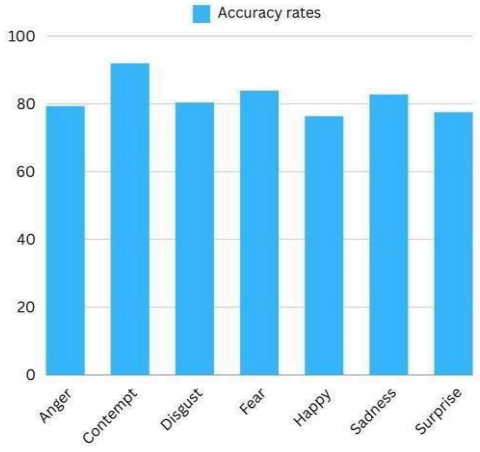


Figure 5: This Figure illustrates the confusion matrix.

These values help evaluate the model's performance and pinpoint areas for development by offering comprehensive information about the model's predictions for each emotion class. True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) can be used to assess accuracy using the following formula: Accuracy = TP+TN+FP+FN / TP+TN

The confusion matrix illustrates where the classification model is making accurate or inaccurate predictions, giving insight into the model's performance. Numerous performance indicators, including accuracy, precision, recall, and F1 score, are frequently computed using it.

Emotion	Precision	Recall	F1 Score	Accuracy
Anger	0.591	0.65	0.62	0.794
Contempt	0	0	N/A	0.920
Disgust	0.63	0.75	0.69	0.804
Fear	0.20	0.20	0.20	0.839
Happy	0.77	0.74	0.76	0.764
Sadness	0.25	0.33	0.29	0.828
Surprise	0.85	0.773	0.78	0.776



Graph 1: This Graph shows the accuracy rates of each emotion.

We can view the graph that displays the accuracy rates of each emotion here. Here, we see that happy has the lowest accuracy rate at 76%, and "contempt" has the highest accuracy rate at 92%.

XIV ..APPLICATION RESTRICTIONS

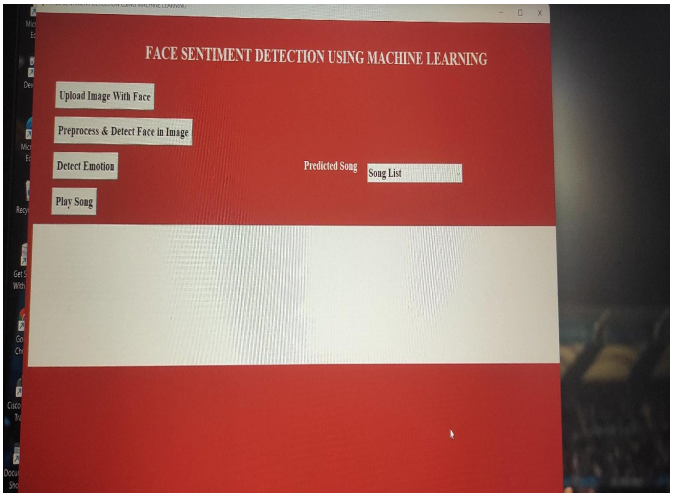
The following is a list of some of the restriction steps that we discovered in our system to aid in the more accurate and efficient operation of software.

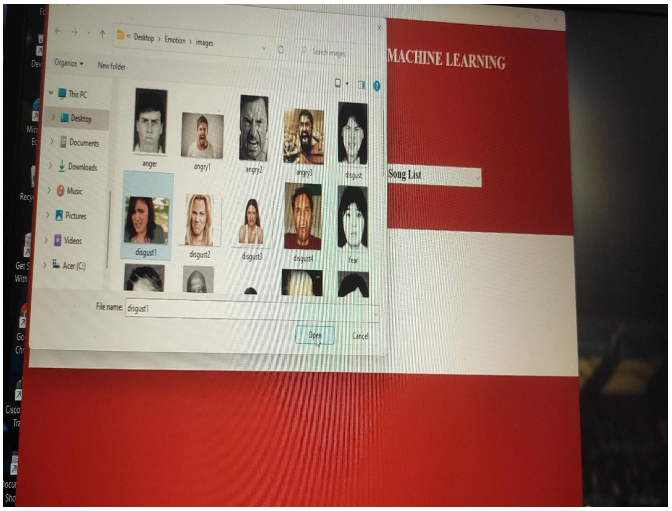
- Real-time recognition becomes quite difficult for photos up to 720p. This is because the face position step execution on the platform takes a very long time. As a result, our system can only identify photos in real time that have a resolution of 180p to 720p.

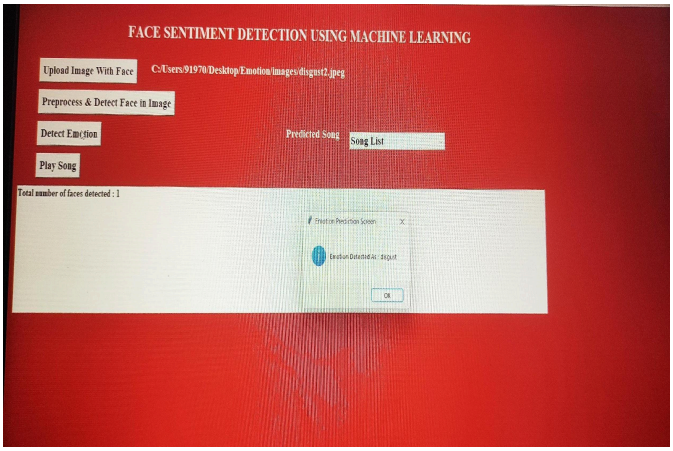
- Face won't be detected if we put the half image i.e. missing face features leads to "no face detected" output.

- Face emotion detection won't be possible with only eyes ,mouth ,or any part of face. It requires entire face and face features and face landmarks to detect feature.

XV . RESULTS







XVI . REFERENCES

1. R. Gupta, S. Tanwar, S. Tyagi and N. Kumar, "Tactile Internet and its applications in 5G era: A comprehensive review", *Int. J. Commun. Syst.*, vol. 32, no. 14, Sep. 2019.
2. Bassili, J. N. (1978). Facial motion in the perception of faces and of emotional expression. *Journal of Experimental Psychology: Human Perception and Performance, 4*(3), 373–379. https://doi.org/10.1037/0096-1523.4.3.373
3. K. Chakraborty, S. Bhattacharyya, and R. Bag, "A survey of sentiment analysis from social media data," IEEE Transactions on Computational Social Systems, vol. 7, no. 2, pp. 450–464, Apr. 2020, doi: 10.1109/TCSS.2019.2956957.
4. J. Xu, Z. Li, F. Huang, C. L, I -and P. S. Yu, "Social image sentiment analysis by exploiting multimodal content and heterogeneous relations", *IEEE Trans. Ind. Informatics*, vol. 17, no. 4, pp. 2974-2982, 2021
5. L. Sirovitch and M. Kirby, "Low-dimensional procedure for the characterization of human faces," Journal of the Optical Society of America A, vol. 2, pp. 519–524, 1987.
6. U.S. Department of Defense, "Facial Recognition Vendor Test, 2000," Available: http://www.dodcounterdrug.com/facialrecognition/FRVT2000/frvt200 0.htm.
7. M. Kirby and L. Sirovich, "Application of the Karhunen- Loève procedure for the characterization of human faces," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, pp. 831-835, Dec. 1990.
8. http://iccv.org/ Islamic Center of Conejo Valley
9. "Real-Time Facial Expression Recognition: A Survey" by S. S. Anjum, et al.
10. "A Review on Facial Expression Recognition Techniques and Databases" by G. Zhao, et al.
11. "Facial Expression Recognition: A Review on Classification Techniques and Databases" by P. Ekman, et al.Packing tubes on tori: An efficient method for low SNR analog error correction | IEEE Conference Publication
12. DEAP: A Database For Emotion Analysis; Using Physiological Signals Sander Koelstra; Christian Muhl; Mohammad Soleymani; Jong-Seok Lee; Ashkan Yazdani; Touradj Ebrahimi; Thierry Pun; Anton Nijholt; Ioannis Patras
13. Coursera: Machine Learning Supervised Learning in

Machine Learning: Regression and Classification (DeepLearning.AI) | Coursera

14. YouTube: DeepLearning.TV https://www.youtube.com/user/DeepLearningTV

---

15. Facial Expression Recognition with Deep Learning GitHub - muxspace/facial_expressions: A set of images for classifying facial expressions
16. Real-time Facial Expression Recognitionhttps://github.com/omar178/Real-time-Facial-Expression-Recognition
17. Facial Expression Recognition using CNN GitHub - amineHorseman/facial-expression-recognition-using-CNN: Deep facial expressions recognition using Opencv and Tensorflow. Recognizing facial expressions from images or camera stream
18. OpenCV: Official Website https://opencv.org/
19. Dlib: Official Website
20. TensorFlow: Official Website Https://www.tensorflow.org/
21. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI): Recent Issues IEEE Transactions on Pattern Analysis and Machine Intelligence
22. IEEE International Conference on Computer Vision (ICCV): Website http://iccv.org/Islamic Center of Conejo Valley
23. IEEE International Conference on Automatic Face and Gesture Recognition (FG Transactions on Affective Computing): Conference website Transactions on Affective Computing Https://fg2024.org/
24. CVPR (Computer Vision and Pattern Recognition) Conference
25. GitHub Repository for Facial Expression Recognition using Deep Learning Real-time Facial Emotion Detection using deep learning
26. Research Paper: "Facial Expression Recognition with Convolutional Neural Networks" [1612.02903] Facial Expression Recognition using Convolutional Neural Networks: State of the Art
27. GitHub Repository for Facial Expression Recognition using OpenCV and Python GitHub - Shantanu/FaceDetect
28. Conference Paper: "Real-Time Facial Expression Recognition with CNN"https://www.researchgate.net/publication/314437044_Real-Time_Facial_Expression_Recognition_with_CNN
29. GitHub Repository for Facial Expression Recognition using Deep Learning and OpenCV GitHub - amineHorseman/facial-expression-recognition-using-CNN: Deep facial expressions recognition using Opencv and Tensorflow.