ISSN: 2321-2152

IJMECE

International Journal of modern electronics and communication engineering

E-Mail
editor.ijmece@gmail.com
editor@ijmece.com

www.ijmece.com

ISSN2321-2152
www.ijmece .com
Vol 8, Issue 3, 2020

HERO: HUMAN EMOTIONS RECOGNITION FOR REALIZING INTELLIGENT INTERNET OF THINGS

Vasanthamma. G, Maltesh Kamatar, Indira
Assoc. Professor, Asst. Professor, Asst. Professor
gvasreddy@gmail.com, maltkpl@gmail.com, indira.raj.06@gmail.com
Department of CSE, Proudhadevaraya Institute of Technology, Abheraj Baldota Rd,
Indiranagar, Hosapete, Karnataka-583225

ABSTRACT:

Facial expression detection has recently grown in popularity due to the ongoing development of artificial intelligence. In the realm of interface technology, emotion detection is crucial. In interaction technology, non-verbal cues account for two thirds of the communication process, whereas vocal cues account for only one third. To identify expressions on people's faces, researchers utilise a technique called facial emotion recognition (FER). Much of a person's emotional state, mental state, and human viewpoint may be conveyed via facial expressions. Using a mix of gender categorisation and age estimate, this article seeks to discover fundamental human emotions. Facial expressions including happiness, sadness, anger, fear, surprise, and neutrality are the building blocks of emotion. An architecture based on You Look Only Once (YOLO) version 2 and a squeezenet architecture is proposed here for a real-time face expression identification system. One such real-time object detection system is the yolo architecture. This is where it finds its real-time application in facial recognition. Anchor boxes are used to capture these pictures accurately. Squeezenet is the second architecture that is utilised for estimating ages and classifying genders. It offers high-level feature extraction and substantial, accurate object recognition, both of which contribute to very fast picture classification and object detection. With their extensive neural network cross-validation and many hidden layers, both designs outperform competing approaches.

Keywords: YOLO, emotion, data set, CNN, sequeezenet.

INTRODUCTION

A face detection includes classifying image into two classes: one with faces (targets), and other containing the background (clutter) which needs to be removed. Commonalities exist between faces, they vary differently in terms of age, skin color and facial expression, this becomes difficult due to this commonality. The further problem is complicated by differing lighting conditions, image qualities and geometries, partial occlusion and disguise is also a possibility. A face detector should be able to detect the presence of any face under different set of lighting conditions in any background condition. The face detection analysis can be broken into two tasks. The first task is a classification task that takes

some random regions of image as input and outputs a binary value of yes or no, indicating if there are any faces present in the image. The other task is the face localization task which is to take an image as input and output the location of any face or faces within that image as some bounding box/boxes with (x, y, width, height). Smart robots can be built by automatic facial expression application. These bots can be used in various applications like interactive games and service center. There are six universal expressions according to Ekman they are fear, disgust surprise, anger, sadness and happiness. Face variances can be observed to recognize these expressions. For example, we can say a person is happy which can be identified as a gesture of smile by tightened eyelids and raised lips corners. A person's internal states, social communication and intentions are indicated by change in facial expressions. Many applications in many areas like human emotions analysis, natural Human computer interaction, image retrieval and talking bots have a large effect on them by automatic facial expression detection. Face Recognition with Histogram of Oriented Gradients using CNN detection has been an impacting issue in the technological community as human beings fined facial expressions one of the most natural and powerful means to express their intentions and emotions. Last stage of the system is facial expression detection. There are basically three steps in training procedure in expression recognition systems named as feature learning, classifier construction and feature selection. Feature learning stage is first, feature selection is second and the last one is classifier construction. Only learned facial expressions variations among all features are extracted after feature learning stage. Facial expression is then represented by the best features which are chosen by feature selection. Not only maximizing inter class variation but they also should try to minimize the intra class variations of expressions not only maximizing inters class variation but they also should minimize the intra class variations of expressions. Because same expressions of different individuals in image are far from each other in pixel's space so minimizing the intra class variation of expressions is a problem. Techniques that can be used for facial detection are YOLO, SDD, RCNN, Faster RCNN.

LITERATURE SURVEY

2.1 Facial Expression Recognition with Histogram of Oriented Gradients using CNN.

Authors: Fayyaz Ali,Sahar Zafar,Irfan Ali,Subhash Guriro,Asif KhanAdnan Zaidi Abstract: A new method is introduced in this study for Facial expression recognition using FER2013 database consisting seven classes consisting (Surprise, Fear, Angry, Neutral, Sad, Disgust, Happy) in past few decades, Exploration of methods to recognize facial expressions have been active research area and many applications have been developed for feature extraction and inference. However, it is still challenging due to the high-intra class variation. Methods/Statistical Analysis: we deeply analyzed the accuracy of both handcrafted and leaned aspects such as HOG. This study proposed two models; (1) FER using Deep Convolutional Neural Network (FER-CNN) and (2) Histogram of oriented Gradients based Deep Convolutional Neural Network (FER-HOGCNN). the training and testing accuracy of FER-CNN model set 98%, 72%, similarly Losses were 0.02, 2.02 respectively. On the other side, the training and testing accuracy of FER-HOGCNN model set 97%, 70%, similarly Losses were 0.04, 2.04. Findings: It has been found that the accuracy of FER-HOGCNN model is good overall but comparatively not better than Simple FER-CNN. In dataset the quality of images are low and small dimensions, for that reason, the HOG loses some important features during training and testing. Application/Improvements: The study helps for improving the FER System in image processing and furthermore, this work shall be extended in future, and order to extract the important features from images by combining LBP and HOG operator using Deep Learning models.

2.2 Active Clustering with Ensembles for Social structure extraction

Authors: J. R. Barr, L. A. Cament, K. W. Bowyer, and P. J. Flynn. **Abstract:** We introduce a method for extracting the social network structure for the persons appearing in a set of video clips. Individuals are unknown, and are not matched against known enrollments. An identity cluster representing an individual is formed by grouping similar-appearing faces from different videos. Each identity cluster is represented by a node in the social network. Two nodes are linked if the faces from their clusters appeared together in one or more video frames. Our approach incorporates a novel active clustering technique to create

more accurate identity clusters based on feedback from the user about ambiguously matched faces. The final output consists of one or more network structures that represent the social group(s), and a list of persons who potentially connect multiple social groups. Our results demonstrate the efficacy of the proposed clustering algorithm and network analysis techniques.

2.3 Fast human detection using a cascade of histograms of oriented gradients

Authors: Q. Zhu, M.-C. Yeh, K.-T. Cheng, and S. Avidan

Abstract:

We integrate the cascade-of-rejectors approach with the Histograms of Oriented Gradients (HoG) features to achieve a fast and accurate human detection system. The features used in our system are HoGs of variable-size blocks that capture salient features of humans automatically. Using AdaBoost for feature selection, we identify the appropriate set of blocks, from a large set of possible blocks. In our system, we use the integral image representation and a rejection cascade which significantly speed up the computation. For a 320 × 280 image, the system can process 5 to 30 frames per second depending on the density in which we scan the image, while maintaining an accuracy level similar to existing methods.

2.4 Local directional ternary pattern for facial expression recognition

Authors: Ryu, A. R. Rivera, J. Kim, and O. Chae

Abstract: This paper presents a new face descriptor, local directional ternary pattern (LDTP), for facial expression recognition. LDTP efficiently encodes information of emotion-related features (ı.e., eyes, eyebrows, upper nose, and mouth) by using the directional information and ternary pattern in order to take advantage of the robustness of edge patterns in the edge region while overcoming weaknesses of edge-based methods in smooth regions. Our proposal, unlike existing histogram-based face description methods that divide the face into several regions and sample the codes uniformly, uses a two-level grid to construct the face descriptor while sampling expression-related information at different scales. We use a coarse grid for stable codes (highly related to non-expression), and a finer one for active codes (highly related to expression). This multi-level approach enables us to do a finer grain description of facial motions while still

ISSN2321-2152
www.ijmece .com
Vol 8, Issue 3, 2020

characterizing the coarse features of the expression. Moreover, we learn the active LDTP codes from the emotion-related facial regions. We tested our method by using person-dependent and independent cross-validation schemes to evaluate the performance. We show that our approaches improve the overall accuracy of facial expression recognition on six data sets.

EXISTING SYSTEM

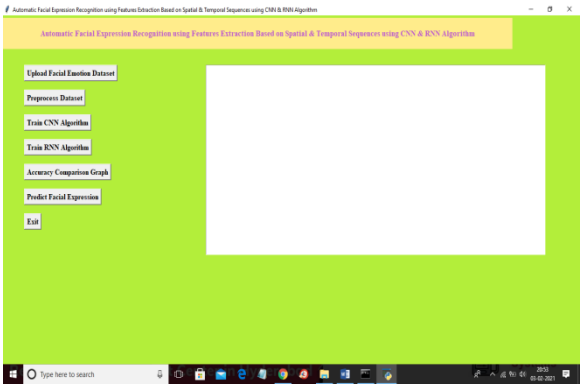
The emotion recognition plays a major role in interaction technology. In interaction technology the verbal components only play a one third of communication and the non-verbal components plays a two third of communication. A facial emotion recognition (FER) method is used for detecting facial expressions. Facial expression plays a major role in expressing what a person feels and it expresses inner feeling and his or her mental situation or human perspective

PROPOSED SYSTEM

This paper aims to identify basic human emotions with the combination of gender classification and age estimation. The facial emotions such as happy, sad, angry, fear, surprised, neutral emotions are considered as basic emotions. Here proposes a real time facial emotion recognition system based on You Look Only Once (YOLO) version 2 architecture and a squeezenet architecture. The yolo architecture is a real time object detection system. Here it used for identify and detect faces in real time. These images are captured by using anchor boxes for accuracy. The second architecture is squeezenet and is used for gender classification and age estimation. It provides significant, accurate object detection and extracts high-level features that help to achieve tremendous performance to classify the image and detecting objects. Both the architectures provide accurate result than other methods with the large no of hidden layers and cross validation in the neural network.

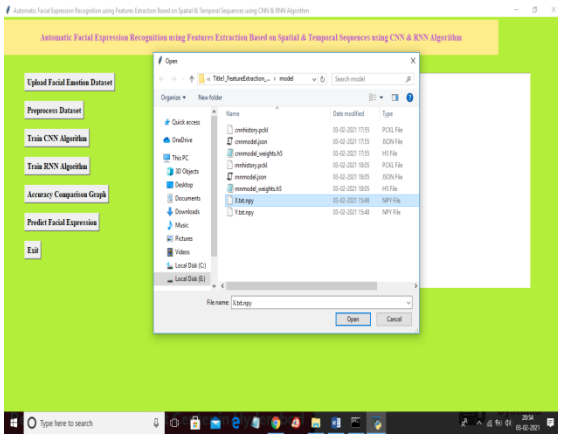
IMPLEMENTATION

In above screen go inside title1 folder and then click on ‘run.bat’ file to get below screen

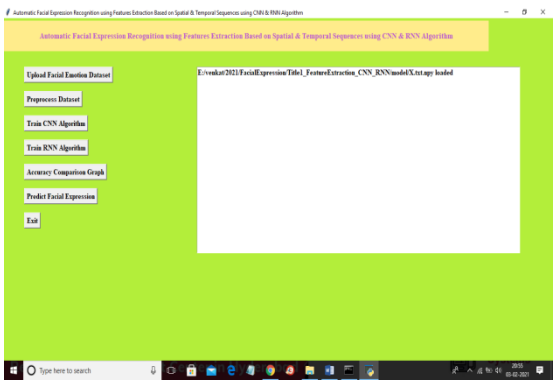


In above screen click on ‘Upload Facial Emotion Dataset’ button to upload dataset

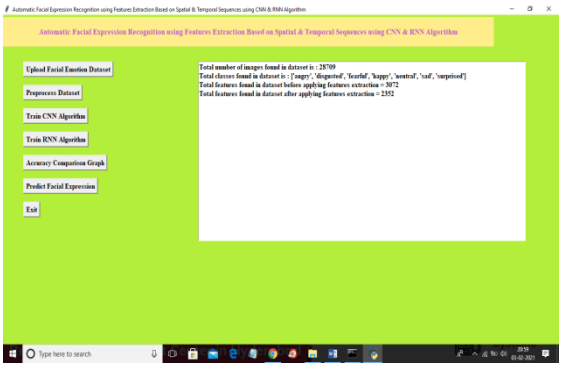
ISSN2321-2152
www.ijmece .com
Vol 8, Issue 3, 2020



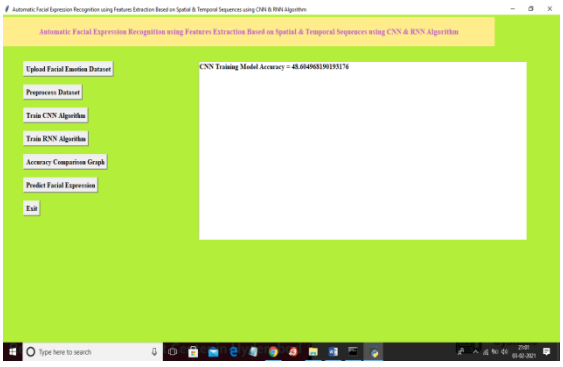
In above screen selecting and uploading 'X.txt.npy' file which contains images of all emotion faces and then click on 'Open' button to load dataset and to get below screen



In above screen dataset loaded and now click on 'Preprocess Dataset' button to read all images and then apply feature extraction algorithm called PCA to read important features from dataset and to get below screen. This module may take 5 to 8 minutes time to give output so please wait till process complete like below screen



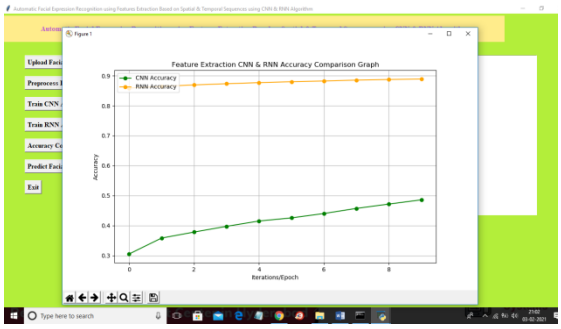
In above screen we can see dataset contains total 28709 images and before applying feature extraction algorithm total images features/pixels are 3072 and then after applying features reduces to 2352 as PCA remove unimportant pixels and used only important pixels/features. Now image data is ready and now click on 'Train CNN Algorithm' button to train CNN with process image features



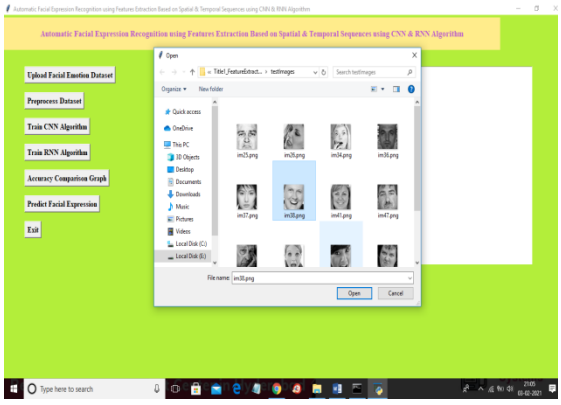
In above screen CNN accuracy is 48 and now click on 'Train RNN Accuracy' button to train dataset with RNN

ISSN2321-2152
www.ijmece .com
Vol 8, Issue 3, 2020

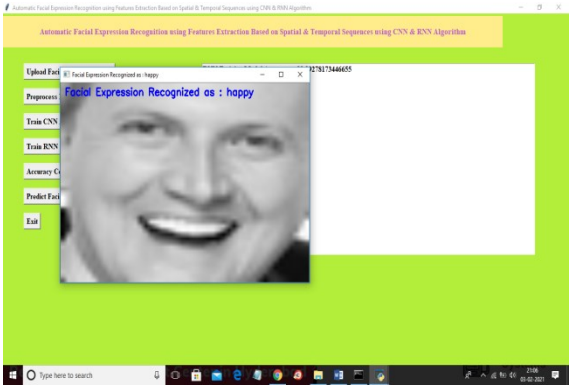
In above screen RNN accuracy is 88% and now click on 'Accuracy Comparison Graph' to get below graph of both algorithms



In above screen x-axis represents epoch/iteration ad y-axis represents accuracy and in above graph orange line represents RNN accuracy and green line represents CNN accuracy and from above graph we can see with further epoch/iteration both algorithm accuracy get better and better and from above graph we can conclude that RNN is giving better result. Now click on 'Predict Facial Expression' button to upload new test image and the application predict emotion from it

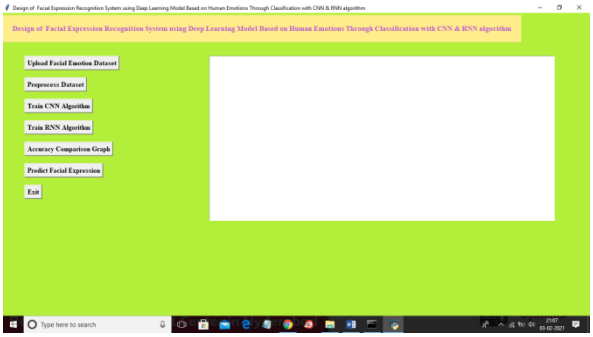


In above screen selecting and uploading im38.png image and then click on 'Open' button to get below result

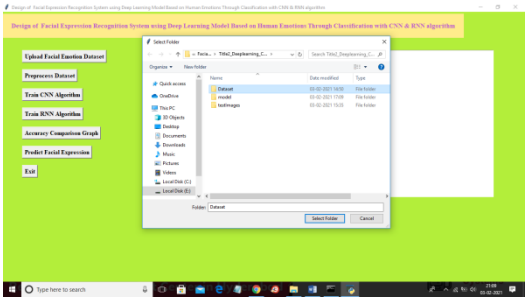


In above screen we got detected emotion as 'happy' and similarly you can upload any image and then predict emotion. So this is the output of TITLE 1.

Now run title 2 project by double click on 'run.bat' file from 'Title2_Deeplearning_CNN_RNN' folder to get below screen



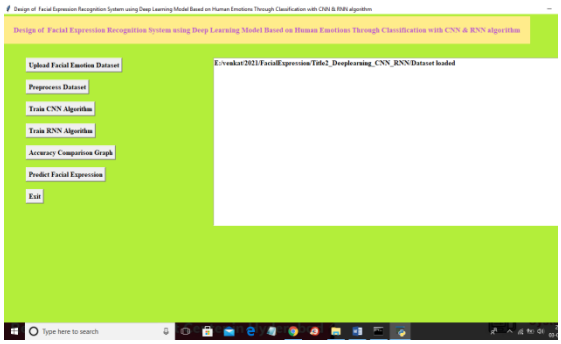
In above screen click on 'Upload Facial Emotion Dataset' button to load dataset and to get below screen



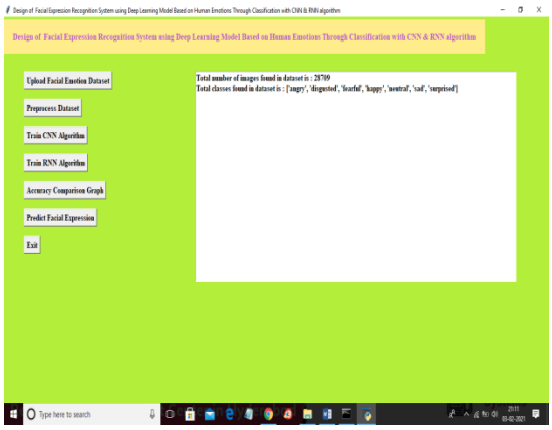
In above screen select and upload 'Dataset' folder and then click on

ISSN2321-2152
www.ijmece .com
Vol 8, Issue 3, 2020

'Select Folder' button to load dataset and to get below screen



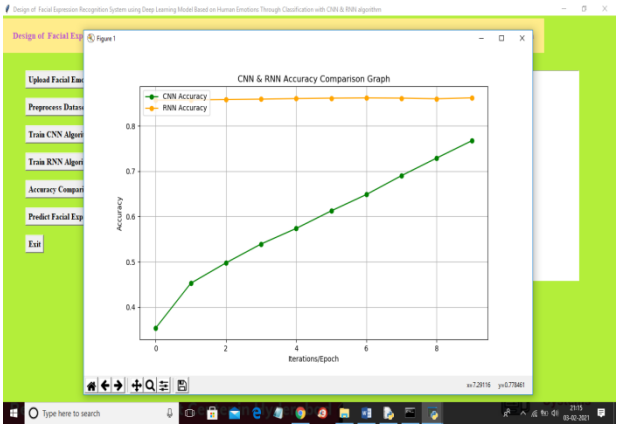
In above screen dataset loaded and then click on 'Preprocess Dataset' button to read all images for training and to get below screen



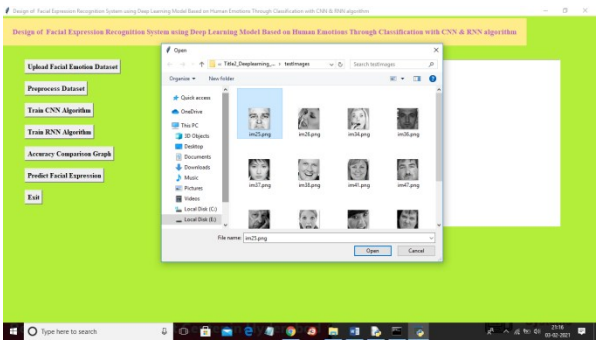
In above screen application read 28709 images from 7 different emotions and now dataset is ready and now click on 'Train CNN Algorithm' button to train above dataset.

In above screen CNN accuracy is 76% and now click on 'Train RNN Algorithm' button to train dataset with RNN algorithm

In above screen RNN train with accuracy as 86% and now click on 'Accuracy Comparison Graph' to get below graph

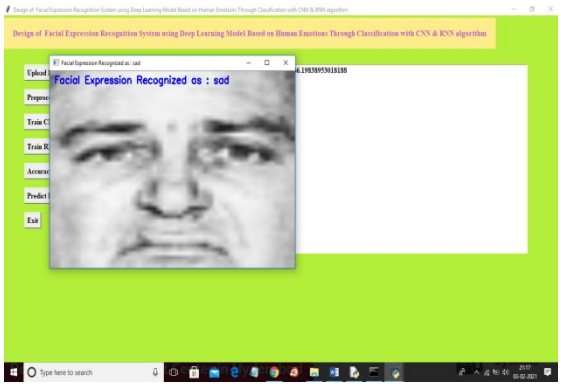


Now click on 'Predict Facial Expression' button to upload test image and then application will predict emotion



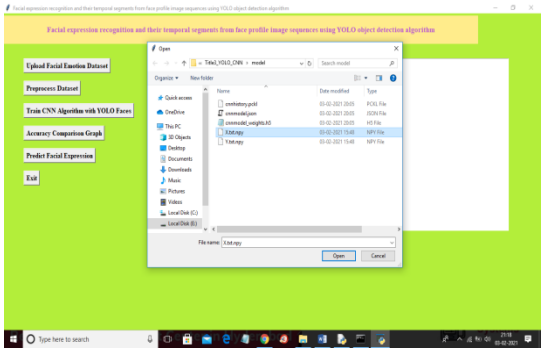
In above screen selecting and uploading im25.png and then click on 'Open' button to get below result

ISSN2321-2152
www.ijmece .com
Vol 8, Issue 3, 2020

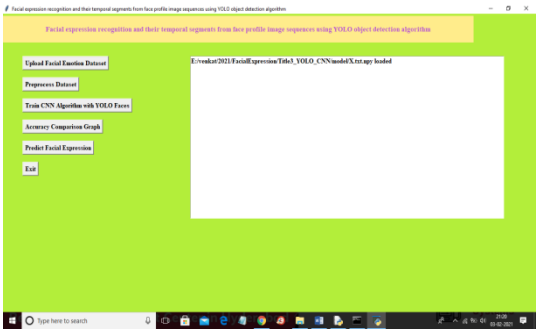


In above screen emotion detected as 'sad'. Now run title3 by clicking on 'run.bat' file from 'Title3_YOLO_CNN' folder to get below screen

In above screen click on 'Upload Facial Emotion dataset' button to load dataset and to get below screen



In above screen select and upload 'X.txt.npy' file which contains all emotion faces and then click on 'Open' button to load dataset and to get below screen



In above screen dataset loaded and similarly you can click on all button to get output and its accuracy details. Similarly you run title4 project and upload dataset and then run all modules to get error rate.

CONCLUSION

In the previous few decades, machine use has skyrocketed. These days, a wide variety of sectors rely on machines. The more time they spend interacting with humans, the more natural and effortless their interactions must become. This can only be accomplished if robots are endowed with the capacity to comprehend their physical surroundings. Especially a person's intents. To this day, computer scientists still face a challenging and complicated task when trying to identify emotions in human expressions. To improve the accuracy and efficiency of real-time face expression detection, this paper proposes a system that combines two deep neural network-based techniques, namely Yolo version 2 and squeezenet architecture. In the future, when feelings are recognised, an action may be taken. Feeling down? The system may lift your spirits by playing a music, telling a joke, or even messaging your closest friend. This might be the next big thing in artificial intelligence—a system that can read the user's emotions and respond appropriately. The divide between people and robots is reduced by this. Additionally, we may include an

interactive keyboard that allows users to just press a button, and the software will detect the user's emotional state and transform it into their preferred emoticon.

REFERANCES

[1] Jumani, S.Z., Ali, F., Guriro, S., Kandhro, I.A., Khan, A. and Zaidi, A., 2019. Facial Expression Recognition with Histogram of Oriented Gradients using CNN. Indian Journal of Science and Technology, 12, p.24.

[2] J. R. Barr, L. A. Cament, K. W. Bowyer, and P. J. Flynn. Active clustering with ensembles for social structure extraction. In Winter Conference on Applications of Computer Vision, pages 969–976, 2014

[3] Ren, S., He, K., Girshick, R. and Sun, J., 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems (pp. 91-99).

[4] Ren, S., He, K., Girshick, R. and Sun, J., 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems (pp. 91-99).R. Goh, L. Liu, X. Liu, and T. Chen. The CMU face in action (FIA) database. In International Conference on Analysis and Modelling of Faces and Gestures, pages 255–263. 2005.

[5] L. Wolf, T. Hassner, and I. Maoz. Face recognition in unconstrained videos with matched background similarity. In Computer Vision and Pattern Recognition, pages 529–534, 2011

[6] Y. Wong, S. Chen, S. Mau, C. Sanderson, and B. C. Lovell. Patchbased probabilistic image quality assessment for face selection and improved video-based face recognition. In Computer Vision and Pattern Recognition Workshops, pages 74–81, 2011.

[7] B. R. Beveridge, P. J. Phillips, D. S. Bolme, B. A. Draper, G. H. Givens, Y. M. Lui, M. N. Teli, H. Zhang, W. T. Scruggs, K. W. Bowyer, P. J. Flynn, and S. Cheng. The challenge of face recognition from digital point-and-shoot cameras. In Biometrics: Theory Applications and Systems, pages 1–8, 2013

[8] N. D. Kalka, B. Maze, J. A. Duncan, K. A. O Connor, S. Elliott, K. Hebert, J. Bryan, and A. K. Jain. IJB–S: IARPA Janus Surveillance Video Benchmark. In IEEE International Conference on Biometrics: Theory, Applications, and Systems, 2018.

[9] M. Singh, S. Nagpal, N. Gupta, S. Gupta, S. Ghosh, R. Singh, and M. Vatsa. Cross-spectral cross-resolution video database for face recognition. In IEEE International Conference on

Biometrics Theory, Applications and Systems, 2016

[10] X. Zhu and D. Ramanan, "Face detection, pose estimation, and landmark localization in the wild," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2012, pp. 2879–2886.