International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-10, August 2019

Artificial Intelligence based Human Facial Action Recognition by Deep Learning Neural Network

Suvarna Nandyal, Lakshmi

Abstract: ***The modern era of digital devices which will be equipped with the dedicated machine will have the built in features to perform the special needs of the human. The mobile devices comes with the special cameras which will produces the high quality images. Human face and the facial expression system has been in the interesting field due to the great level of application it possesses. The face unlock, face detection in security areas and also the facial expression recognition has been in demand of development. The human exhibits the multiple type of expressions anger, sad, surprise, happy and many more. These expressions have special notation based on the human feelings and scenario. The Digital image processing technology has provided the tools to work with the images and will use the edge features, skin mapping and supervised knowledge based classifier to recognize the expressions posed by the human to assist the need.***

Index Terms: Face, GLCM(Gray Level Co-occurrence Matrix), Skin mapping, SVM(Support Vector Machine).

I. INTRODUCTION

Survivalance systems has been in the trending as the security of any human or individual is essential in the modern life. This is basic human behavior essential for effective communication and interaction among people, such as face recognition, safety monitoring and surveillance systems.

Globally terrorist activities are increasing day by day, to identify such people in public places is difficult. If we can succeed in identification of correct emotional state of the human being and then train machine system to identify expressions, then probably terrorist will be identified easily even at the crowded place. To determine emotional state of any individual it is required to know his or her facial expressions. Facial expression is a bodily state involving various physical structures; it is gross or fine grained behavior, and it occurs in a particular situations. Any empirical research on facial expression involves only some part of the broad meaning of them. After over a century of years terrorists activities are still existing. Identifying such terrorists depends upon what kind of facial expressions and how they are communicated. Facial expressions are related to external physical expressions. His/her external body expression; It also includes internal feelings and thoughts as well as other internal processes that the person may not be aware of. In world wide several facial expression recognition systems, methods and algorithms have been proposed for facial analysis from both static images and image sequences. Surviavalance systems has been in the trending as the security of any human or individual is essential in the modern life. The devices which we use for security like automatic door, lockers, signal systems and many has to be triggered before it can perform the actual work. As we know for any human the expression itself will describe the state of the mind of the person. It will describe the human inner feelings also. The process of extracting the face, recognizing the expression based on the Geometrical features on the face will be a challenging aspect, as many user poses different type of the expression.

Manuscript published on 30 August 2019.
***Correspondence Author(s)**
Dr.Suvarna Nandyal , Computer Science and Engineering, PDACE, Kalaburgi, India, suvarna_nandyal@yahoo.com.
Lakshmi, Computer Science and Engineering, PDACE, Kalaburgi, India, lakshmissh02@gmail.com.

II. LITERATURE REVIEW

Over the past few years, several more methods have been proposed. Each system has used a special Classification methods, filters to recognize the expressions. A few of them are discussed in this section.

Ping Liu et al.(2014) proposed a Boosted Deep Belief Network (BDBN) in the framework of multiple layered unified loopy. A training process for the face expression has been performed in the three individual stages: feature learning, selection, and the feature classifier based data construction. These features will be trained and features are extracted in the extreme manner by using the dataset of the initial images**.** By using the BDBN framework, the feature set, generally used to classify the main characteristics of the expression of human facial related appearance and also the shape changes, boosted strong classifier by these feature set in the statistical way.

Yachen Zhu et al.(2014) proposed the multiple human facial action unit based expression recognition by the dataset modeling . the image features and the image target labels are used to classify the images. Initially the multi-task image dataset based feature learning has been adopted for the purpose of dividing the action unit recognition job into the multiple groups, later each of the shared features for every single group. Later, the Bayesian network has been used by the author to co-existent and also the data based mutual exclusive semantic in between the target labels of images of input. **Adria Ruiz et al.(2015)** In this paper the author has investigated the way of using the large database, which is labeled based on the 6 universal facial expressions. A novel training based human facial expression recognition has been proposed by the author based by using the Hidden-Task Learning.

HTL exploits the relationship between hidden and visible tasks.. The values based dataset will be analyzed and classified based on the trained dataset, the use of these values will help in accurate classification of human face expression.
Wei-Yun Yau et al.(2015) Natural human image based emotion recognition has been proposed by the author as it has the wide range applications , like the human-computer interaction, tutoring systems, smart devices, driver assistance etc. generally the process is carried out by using the laboratory controlled image based data, which may not be exact reflection of the environment in real time. For dynamic recognize of the human facial emotions in real time based natural situations, the method known as the Extreme Sparse Learning (ESL), which is capable of creating the local dataset and perform the classification based on the previous history.
Kaili Zhao et al. (2015) proposed FACS based human emotion detection system. The Facial Action Coding System (FACS) has been adopted by many of the institutions as it possesses many of the possible facial location based on the human face shape. set of the dataset which has been given will help the user to resolve the complexity in analyzing the data. Visible faces are subdivided according to the effect of facial muscles in more than 30 behavior units (AU).. These combinations will be divided into number of combinations again to get the accurate result in real time.
Gil Levi et al. (2015) presented a novel Classifying emotions by using the static facial images has been proposed by the author. By adopting the success of the Convolutional Neural Networks (CNN) on the human facial recognition problems. Less labeled data features from the input image which is available for the training the human face emotion classification systems. It deals with the illumination factors, edge based factors. The proposed method will be able to perform the accurate image classification.
Stefanos Eleftheriadis et al. (2015) proposed a new multi-conditional sub population model for simultaneous face fusion and detection of facial activity units. This approach takes advantage of the differentiating features of existing filters, such as the structure discovery and logistics functions of generation models such as Gaussian processes. This superior performance result is compared to the existing classifier for the target task using differential or generative attributes, but not for both.
Mina Bishay et al. (2017) proposed a novel Deep Learning architecture that fuses information from several specialized Deep Neural Networks(DNNs) each of which models a different aspect of the problem in question. At the core of this approach is a novel dynamic adaptation of the Deep Network cost function so as to deal with the data imbalances that are inherent in multi label classification problem this allows cross database training. The benefits of the proposed training approach and how different architectures are more suitable for particular AUs. Extensive experimental results show that multi-modal approach outperform the state of the art by a considerable margin.

III. PROPOSED WORK

The proposed system will have the following methodologies:

- Image acquisition by using the system storage. Collection System Use the storage device to collect images
- The image is then passed to the system for the processing where the image is processed at multiple factors.
- The processes are the edge hole filled, contrast enhancement, face part extraction by skin mapping.
- Then cut out the extracted face parts and save them to the system storage for later use.
- The face image is then used for the extraction of the face parts like eyes and the mouth part.
- The extracted face parts are proceeding in graphical phase with the SVM classification to give the expression recognized.

IV. DESIGN AND IMPLEMENTATION

System Block diagram

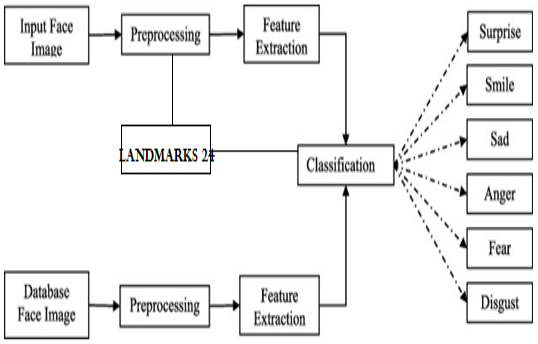

Figure 1: system block diagram

Implementation

Several methods are used to locate, detect, extract the human face also classify the expression based on the previous data.
A. Image acquisition: When acquiring an image, recording text acquires the scanned image as an associated input image and should be in specific format like JPEG, BMT etc. These pictures cannot be inherited by cameras, scanner or the other applicable digital data entry devices. This command reads the file's grayscale from the file by the string computer filename. If the file is not found in existing folder or within the MATLAB path, then it specifies the total path name. In the initial step of any image processing based work, the given input image is processed before passing it to the main system. The preprocessing of the image includes image resizing, contrast adjustment, brightness adjustment, image cropping, image rotation etc. The output of the preprocessing will be the lab image which will be suitable for next processing.

- **Image Resizing:** Reducing the overall size of the image by [height, width] ratio. It is required to increase or decrease the total number of pixels, zooming refers to increase the quantity of pixels, so zooming an image will give more detail.
  Img = imread ("Filename.jpg");
  //img is the original image of 1024*1600 pixels
  Resizeimage=imresize (img, [255,255];
  //Hence the image is resized and reduced to 255*255

- **Image Restoration:** The process of importing corrupted images and sharpening them and estimating the original image. Corruption can be caused by motion blur, noise and camera misfocus. The goal of image restoration technology is to reduce noise and recover resolution loss. The restoration process allows the user to adjust the contrast, brightness and the other features of the image. This is not a noise reduction method, but allows the user to smoothening the image by using the simple method like point spread function and de convolution method etc.
- **Image De-noising:** Noise is an unwanted signal in the image. This may be due to memory transmission or processing. Images can be deleted from some noise image inputs, such as Gaussian noise, impulse noise, coherent noise, granular noise, poison noise, sub noise, toxic noise etc. It is a process of removing the possible noise which may be present in the image input. The removing of noise from the image is called Image De-noising. Images will be get corrupted due to the transmission errors, storage time or addition of unwanted noise to the image. Image de-noising has been a major issue in digital image processing for many years.

Many of the techniques have been proposed like image de-noising with Wavelet transformation, Fuzzy logic transformation, and threshold based de-noising, Spread spectrum, Multilevel clustering of the pixels etc.

B. Human face processing: At this stage of the human face processing, initially processed face will be transformed based on the requirement in the phases like Lighting compensation, Edge based Unit analysis, Edge hole filling. By these extracted feature set of type human face, the skin mapping is performed by using the following module.

C. Face part extraction (skin mapping): Skin mapping plays an important role because it is used as the basis for subsequent steps in image extraction based on extracted pixel values and extraction. separates entities by these labelled pixels. Thus, the probability of an image pixel of the form $c = [Cb\ Cr]^T$, is

$$p(c/skin) = \sum_{j=1}^{M} p(c/j).P(j) \qquad (.1)$$

,*M*- number of Gaussian components model

P(j), the probability pixel at jth location which is ,

$$p(c/j) = \frac{\exp\left[-\frac{1}{2}(c-\mu_j)^T \Sigma_j^{-1}(c-\mu_j)\right]}{2\pi\sqrt{|\Sigma_j|}}$$

μ_j- mean

Σ_j (.2)variance matrix of the j^{th} pixel

In this case, pixel mapping to the skin of the image is important. This is because the input image can be take into account unwanted objects in the path. Skin pixel mapping by GLCM is important to resolve and remove unwanted parts. The Gray level co-occurrence matrix is used to extract pixels from the skin of the image. This matrix is a displacement vector(d), represented as P[i, j] and grouped by pixel value[3*3]. Each pixel value in the original image are copied and replaced by the equitant image pixel in gay level scale.

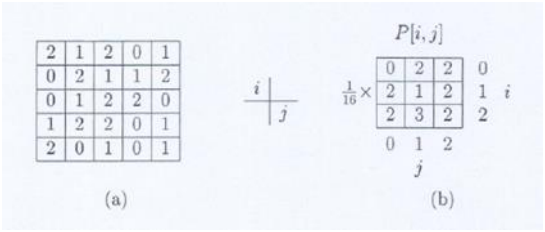

Fig 2: GLCM processing pixel by pixel

D. GPA based SVM classifier: SVM(Support vector machine) is a supervised machine learning method that inspects data and identifies similar model types used in subsequent classification. SVM model was able to model the complex structure of the non-linear decision boundaries with high accuracy and SVM is efficiently used for the binary classification. Landmark-24 is a universally accepted recycled dataset once facial and facial image analysis is complete. A set of molecular data characteristics (performed in the system) will be used to classify the type of component detected. The categorizer uses the extracted feature set to rank the closest matching geometry coordinates.

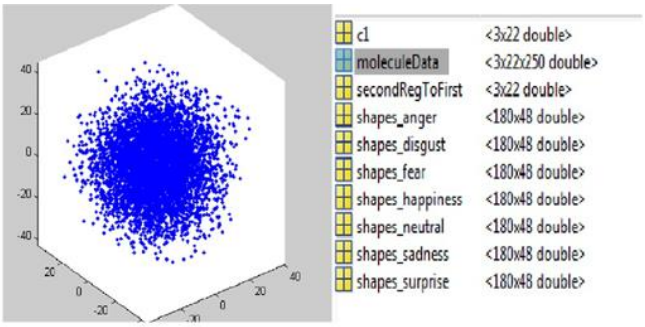

Figure 3 : GPA feature set pretrained

The main advantage of the algorithm is that it categorizes the types of input requests based on feature-based vectors and learning samples. Configured features and test images are supported. Trained features and the testing images are than classified using the Support Vector Machine classifier to get the result to the user .

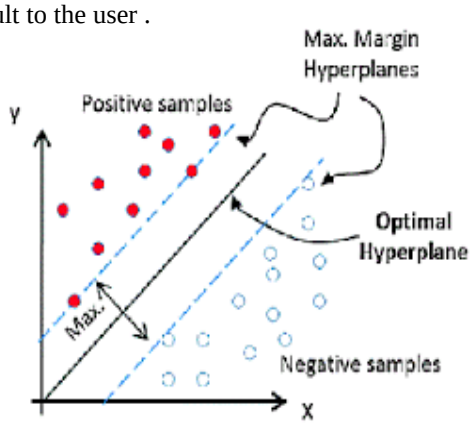

Figure 4 : General SVM classifier working

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

As shown in the above diagram, the SVM classifier extract the patterns from the input sample in our case it is a image. It will be classify the image sample based on the type of the input. The SVM is composed of the followings:

- Hyper planes: Hyper planes with optimal maximum minimum margins are used for pattern classification
- Negative samples: which will drop below the required threshold values
- Positive samples: which will drop above the threshold value

The human face is treated as an input image This is due to the image processing method that extracts only a portion of the face and processes it for use by the user..

E. Face Extraction Algorithm

Algorithm Steps:

Step 1: Iitialization
Step 2: Give input image to the system from the data set.
Step 3: Perform the basic image processing operation from the given input image.
Step 4: Classify images as Training and Testing data set.
Step 5: Pass the processed image set to the system from Feature Extraction and image labelling.
Step 6: Based on the features and results classify images by SVM.
Step 7: Interpret the obtained result to the user as a require.
Step 8: Stop the system.

V. RESULT AND DISCUSSION

The system takes the face image as an input, it performs the several operations like lighting compensation, edge hole filled, image edge dilation and so on and identifies the particular face image for the further process. The eyes and mouth parts are extracted. Then the SVM classifier is applied in order to identify the expression, it is given with a three inputs like testing image, trained image and the classifications. The dataset consisting of pretrained face images that exhibit different expressions. The testing image is compared with pretrained image to identify the expression expressed by that human. The result is shown using MATLAB(Matrix Laboratory). It acts like a laboratory and is very easy language because it has ‘major code part.

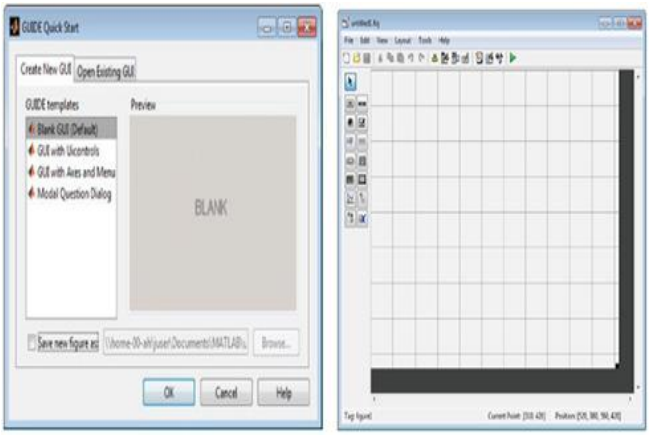

Figure 5: Guide and component pallet

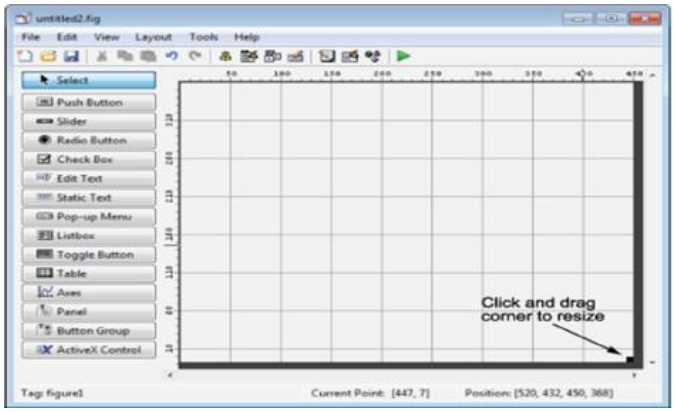

Figure 6: Resizing grid in layout editor

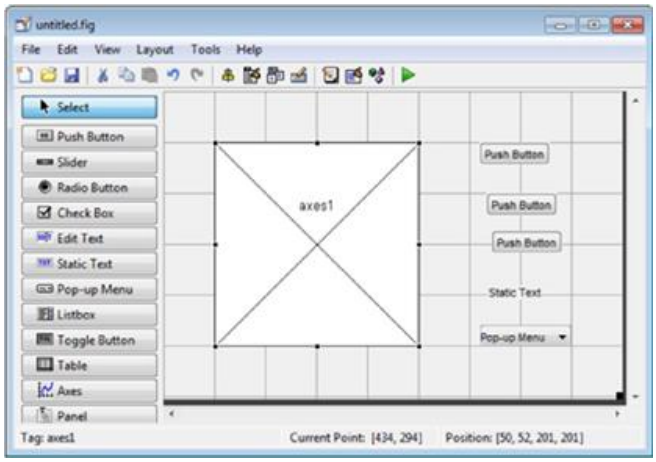

Figure 7: Add axes, buttons and statics texts

Use This Tool...	To...
	Components to the GUIDE Layout Area" on page 6-17 for more information.
Figure Resize Tab	Set the size at which the UI is initially displayed when you run it. See "Set the UI Window Size in GUIDE" on page 6-11 for more information.
Menu Editor	Create menus and context, i.e., pop-up, menus. See "Create Menus for GUIDE UIs" on page 6-91 for more information.
Align Objects	Align and distribute groups of components. Grids and rulers also enable you to align components on a grid with an optional snap-to-grid capability. See "Align GUIDE UI Components" on page 6-79 for more information.
Tab Order Editor	Set the tab and stacking order of the components in your layout. See "Customize Tabbing Behavior in a GUIDE UI" on page 6-88 for more information.
Toolbar Editor	Create Toolbars containing predefined and custom push buttons and toggle buttons. See "Create Toolbars for GUIDE UIs" on page 6-108 for more information.
Icon Editor	Create and modify icons for tools in a toolbar. See "Create Toolbars for GUIDE UIs" on page 6-108 for more information.
Property Inspector	Set the properties of the components in your layout. It provides a list of all the properties you can set and displays their current values.
Object Browser	Display a hierarchical list of the objects in the UI. See "View the GUIDE Object Hierarchy" on page 6-119 for more information.
Run	Save the UI and run the program.
Editor	Display, in your default editor, the code file associated with the UI. See "Files Generated by GUIDE" on page 2-24 for more information.
Position Readouts	Continuously display the mouse cursor position and the positions of selected objects

Figure 8: Guide Tools summary of the MATLAB

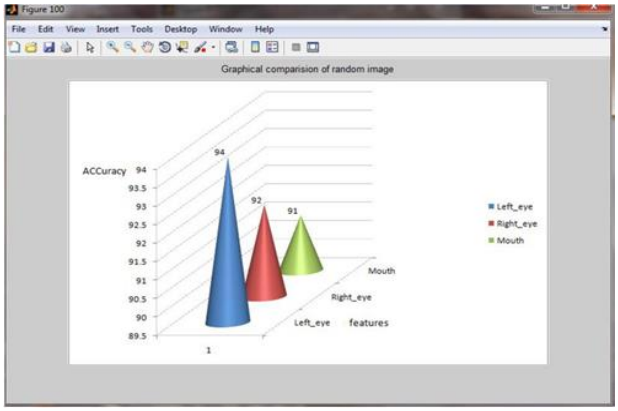

Figure 9 : Graphical Comparison study of different expressions

VI. CONCLUSION AND FUTURE WORK

By this proposed work we are capable of determining the possible nearest expression in the input image. The DIP technology with skin mapping and skin part extractions with the Landmark-24 pre trained dataset has been used to detect and classify the expression. The image is processed initially to extract the feature set and it is later compared to the pre trained dataset. The proposed work has found efficient and accurate to the preset dataset. The values of accuracy with SVM is around 94 to 95%. In future the video enabled face expression tracking at time gap has to be developed to make the system real time application which will need a dedicated camera, device and processor.

REFERENCES

1. Adria Ruiz, Joost Van de Weijer and Xavier Binefa, “From Emotions to Action Units with Hidden and Semi-Hidden-Task Learning”, 5th IEEE International Conference on Computer Vision, pp. 1550-5499, 2015.
2. Gil Levi and Tal Hassner. “Emotion recognition in the wild via convolutional neural networks and mapped binary patterns”, In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction,” pages 503–510. ACM, 2015.
3. Kaili Zhao, Wen-Sheng Chu, “Joint Patch and Multi-label Learning for Facial Action Unit Detection”, p 4673-6964, 2015.
4. Mina Bishay and Ioannis Patras, “Fusing Multilabel Deep Networks for Facial Action Unit Detection,” IEEE 12th International Conference on Automatic Face & Gesture Recognition, vol.17, no.25, pp.5090-5094, 2017.
5. Ping Liu, Shizhong Han, Zibo Meng , and Yan Tong, “Facial Expression Recognition via a Boosted Deep Belief Network”, IEEE Conference on Computer Vision and Pattern Recognition, pp. 1063-1069 ,2014.
6. Seyedehsamaneh Shojaeilangari, Wei-Yun Yau, Karthik Nandakumar, Jun Li, and Eam Khwang Teoh. “Robust representation and recognition of facial emotions using extreme sparse learning”, IEEE Transactions on Image Processing, 24(7):2140–2152, 2015.
7. Stefanos Eleftheriadis and Ognjen Rudovic “Multi-conditional Latent Variable Model for Joint Facial Action Unit Detection” p 4889-4899, 2015.
8. Yachen Zhu, Shangfei Wang, “Multiple-Facial Action Unit Recognition by Shared Feature Learning and Semantic Relation Modeling”, 22nd International Conference on Pattern Recognition, pp.1051-4651, 2014.

AUTHORS PROFILE

Professor Dr.Suvarna Nandyal is professor and HOD of computer science and engineering at Poojya Doddappa Appa College of Engineering in Kalaburgi, Karnataka, India. She holds a bachelor's degree in computer science and engineering from Gulbarga University , Gulbarga in 1993 and a master's degree in computer Science and engineering from VTU Belgaum in 2003. PhD in Computer Science and Engineering at Jawaharlal Technological University in Hydraabad, AndraPradesh, India, in 2013, She undertook a wide range of research in image processing, video processing, and speech, Cloud computing and information retrieval. She is a member of ISTE and IETE. She has published several articles in international and national journals and conferences. She is involved in a variety of academic activities such as program development and professional social activities.

Lakshmi is a PG student in Computer Science and Engineering at Poojya Doddappa Appa College of Engineering, Kalaburgi. Graduated from Godutai Engineering College for Women, Kalaburgi with BE degree in Information Science and Engineering Department in the year 2016. Her research interest in the area of Artificial Intelligence and Machine Learning.