

Patient Monitoring System from Facial Expressions using Python

U.G. Scholar, A. Sudharsan, U.G. Scholar A. Yaswanth, Assistant Professor Mrs.K.Hemapriya

Department of Computer Science and Engineering
Panimalar Institute of Technology, Poonamallee, Chennai
sudharsan2410@gmail.com , yashwan98@gmail.com , hemu.june3@gmail.com

Abstract – In human communication verbal, vocal and non-verbal signals are used to communicate with one another. Usually people tend to precise their emotions either by verbally or vocally, except for those who cannot communicate verbally or vocally, non-verbal signals (Facial Expression) are the only form of communication, which helps to spot the emotions of the person. Basically human facial expressions change with different states of health. Abnormal conditions such as pain is assessed usually through self-reports or by an observer. Reports of pain may be vulnerable to errors due to subjective biases of observers. Moreover, continuous monitoring by humans is impractical. Therefore, automatic pain detection technique have deployed to assist human caregivers and complement their service, thereby improving the standard management, especially for non-communicative patients. Facial expressions are reliable indicator of pain, and are employed in observer-based pain assessment tools. The Remote-target person is monitored by using a webcam to spot the physical condition. The Facial images of the person are captured and analyzed using machine learning. This project proposes a system for recognition of facial expressions using the captured images. This proposed system uses a CNN classifier to differentiate the acquired image into different emotion categories (Happy, Sad, Pain and Normal). Results show that reliable facial expression recognition. If it recognize the person in abnormal condition then alert will be pass to the nearby hospital.

Keywords – Facial expression mapping, Image recognition, Convolutional Neural Network.

I. INTRODUCTION

The field of face expression synthesis has made significant progress. Successful models [1]–[4] are developed to transfer facial expressions of emotions acted by an actor (source) to the face image of the target person. Such models are expensive as an actor is involved. Such models often assume that a large sort of facial expressions of the target person is available.

Collecting this set may not be always feasible. There also exist synthesis models [5], [6] that needs only one expression-neutral face image of the target person. Such methods generally synthesize expressions of basic emotions. But in real life, faces show mix of basic emotions. For example, one may be happily or fearfully surprised. Therefore, we need a model that (a) can synthesize realistic expressions representing both basic as well as mixed emotions and (b) requires only one face image of the target person as input. In this paper we present such a model. In [7] we introduced Expression Map (XM). The XM was shown to be useful for estimating the share of various basic emotions in a given face expression. For example, a facial expression may show 40% happiness and 60% surprise. The current proposed method utilizes the XM to synthesize a specified combination of basic emotional expressions on a given

expression-neutral face image. The AMD features, as used in [7], being histogram in nature are not suitable for synthesis. For synthesis, we use a special set of features from [8] representing facial shape, texture, face structure and expression intensity. The major difference between the features utilized in [8] and that is in the proposed method the texture feature is split into low and high frequency components which are processed separately. This helps us in reducing the noise within the synthesized facial expressions. Using these features we train the XM following the leducational process of Self Organizing Map (SOM) [7], [9]. We visualize expression as a change in facial appearance (shape and texture) from the expression-neutral face. This changes gets stored as a pattern in the node of the trained XM. We propose an objective function that chooses a specific pattern from a node of the trained XM given an emotion (to be synthesized) and an expression-neutral target face image. For synthesizing expression representing the mix of two basic emotions, our method takes weighted combination of two patterns represented by the two nodes of the XM. Thus, the expressive face, synthesized by adding the chosen combination of patterns to the expression-neutral target face image, looks realistic and preserves subject identity. Fig. 1 presents examples of basic and mixed expressions synthesized by the proposed approach. Observe (Fig. 1) that the synthesized expression looks natural on the target face. Different steps of synthesizing

expression are depicted in the form of block diagram in Fig. 2. Our main contributions in this paper are as described. We propose an algorithm that, using the XM, can synthesize basic and mixed (pre-specified combinations of basic emotions) emotional expressions on a given expression-neutral face image. The synthesized expressions are rich in details such as the appearance/disappearance of wrinkles, furrows, teeth etc. and look natural on the facial structure of the target face. In deep learning, a convolutional neural network (CNN or ConvNet) may be a category of deep neural networks, most typically applied to analyzing visual imagery. They are also called as shift invariant or space invariant artificial neural networks, supported on their shared-weights architecture and translation invariance characteristics. They have applications in image and video recognition, recommender systems, image classification, medical image analysis, natural language processing, and financial time series.



Fig. 1. Left: Expression-neutral target face, middle: Synthesized face displaying basic emotion happiness, right: Synthesized face displaying mixed emotion disgust with happiness. The arrows point to the changes in appearance due to imposition of expressions.

The proposed algorithm needs only one expression-neutral face image of the target person. Related works are presented in the next section. Section III describes the feature extraction procedure. The process of expression synthesis is described in Section IV. Section V presents the experimental results. The proposed method is compared with the state-of-the-art works in the same section. The conclusions are summarized in

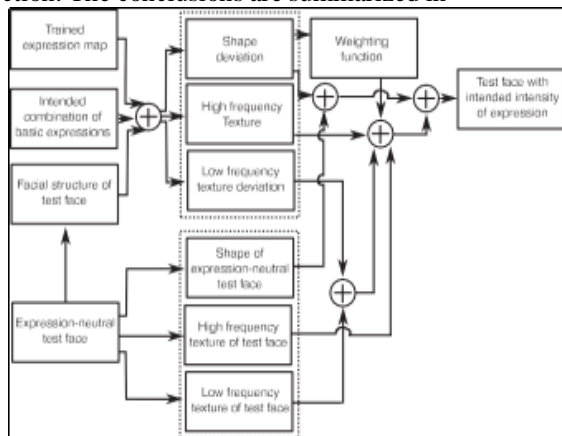


Fig. 2. Block diagram depicting different steps of synthesizing expression on a given expression-neutral face image using the trained expression map.

II. LITERATURE SURVEY

Individual Stable Space: An Approach to face recognition Under Uncontrolled Conditions

There usually exist many kinds of variations in face images taken under uncontrolled conditions, such as changes of pose, illumination, expression, etc. Most previous works on face recognition (FR) focus on particular variations and usually assume the absence of others. Instead of such a "divide and conquer" strategy, this paper attempts to directly address face recognition under uncontrolled Conditions. The key is the individual stable space (ISS), which only expresses personal characteristics. A neural network named ISNN is proposed to map a raw face image into the ISS. After that, three ISS-based algorithms are designed for FR under uncontrolled conditions. There are no restrictions for the images fed into these algorithms. Moreover, unlike many other FR techniques, they do not require any extra training information, such as the view angle. These advantages make them practical to implement under uncontrolled conditions. The proposed algorithms are tested on three large face databases with vast variations and achieve superior performance compared with other 12 existing FR techniques

Discriminate analysis of principal components for face recognition

In this paper we describe a face recognition method based on PCA (Principal Component Analysis) and LDA (Linear Discriminant Analysis). The method consists of two steps: first we project the face image from the original vector space to a face subspace via PCA, second we use LDA to obtain a best linear classifier. The basic idea of combining PCA and LDA is to improve the generalization capability of LDA when only few samples per class are available. Using PCA, we are able to construct a face subspace in which we apply LDA to perform classification. Using FERET dataset we demonstrate a significant improvement when principal components rather than original images are fed to the LDA classifier. The hybrid classifier using PCA and LDA provides a useful framework for other image recognition tasks as well.

Towards Personalizing Classical Music Recommendation

While fans of classical music were found to be underrepresented on social media and music streaming platforms, they constitute an important target group for music recommender systems. We therefore focus on this group of listeners and investigate a wide range of recommendation approaches and variants for the task of music artist recommendation. Within the group of classical music listeners, we further assess categorizing users according to demographics and temporal music consumption behavior. We report the results of preliminary recommendation experiments and insights gained for the listener group under consideration.

Personalized music recommendation algorithm based on tag information. With the development of the mobile Internet and streaming media technology, digital music has been accepted by many people. However, in the face of the massive music data on the Internet, users want to find favorite music is like looking for a needle in a haystack via traditional music indexing techniques. In this paper, an improved algorithm for personalized music recommendation based on tag information is proposed. Firstly, we use the improved user-based collaborative filtering algorithm to deal with the user's long-term preferences. Secondly, according to the user-tag-music relationships, getting the music that associated with the user via recommendation algorithm based on bipartite graph. Experiments on real dataset demonstrate the efficiency and effectiveness of the improved algorithm, getting better results and have practical significance. Music personalized Recommendation System Based on Hybrid Filtration

Due to the diversity and fuzziness of music itself and the accuracy of music melody, the recommendation algorithm with highest accuracy cannot match users' test completely. For such problem, this article proposes a hybrid recommendation algorithm based on collaborative filtering algorithm and music genes, and design a personalized music recommendation system. The system first computes recommendation results according to the collaborative filtering algorithm, and discover the potential interests of the user. Then each music is weighted by preference on the genes of collected music. After weight selection the music which has prior preference is taken as recommendation results. Finally two recommendation results are performed weighted combination and filtering to make recommendation. The experimental data indicate the improved scheme can increase the accuracy of recommendation and meet users' demand from different levels.

III. IMPLEMENTATION

1. Input Video:

The live video taken from the camera is taken as the input video. These videos are captured directly from the person surviving area is taken as input for the further process.

2. Frame Separation:

Frame processing is the initial step within the background subtraction algorithm, the aim of this step is to make the modified video frames by removing noise and unwanted objects within the frame in order to extend the quantity of information gained from the frame. Preprocessing of an image may be a process of collecting simple image processing tasks that change the raw input video info into a format. This could be processed by subsequent steps. Preprocessing of the video is necessary to enhance the detection of moving object's, For example; by spatial and temporal smoothing, snow as moving leaves on a tree,

could be removed by morpholog processing of the frames after the identification of the moving object.

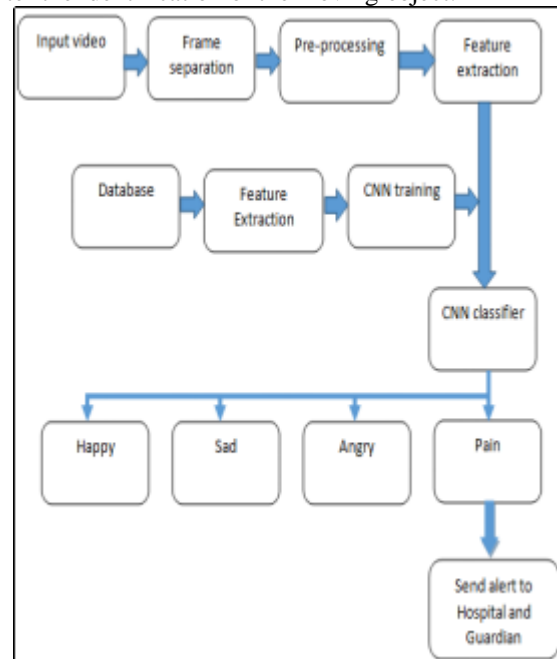


Fig .3. Work flow diagram.

Image pre-processing:

Pre processing is widely used to remove noise from an signaled images . Noise reduction is an important pre-processing step to enhance the results of later processing. The major important role of the median filter is to perform through the signal entry by entry, restoring each entry with the median of neighboring entries.

Image processing mainly include the following steps:

- Importing the image via image acquisition tools;
- Analyzing and manipulating the image;
- Output in which result can be altered image or a report
- which is based on analyzing that image.

3. Feature Extraction:

Feature extraction a form of dimensionality reduction that efficiently represents interesting parts of an image as a compact feature vector. When the input file to an algorithm is too large to be processed and it is suspected to be redundant then it could be transformed into a reduced set of features. Determining a subset of the initial features is termed feature selection. The selected features are expected to contain the relevant information from the input file, so that the required task can be performed by using this reduced representation instead of the entire initial data. This approach is useful when image sizes are large and a reduced feature representation is required to quickly complete tasks like a image matching and retrieval.

4. Database:

The database contains the pre-defined face patterns from feature extraction with which the user's face is compared

and emotions are detected. Pre-defined face patterns are used to compare with the input images.

5. CNN Algorithm:

Artificial Neural Networks are utilized in various classification task like image, audio, words. To get best results using the neural network, it is necessary to settle on an appropriate architecture and learning algorithm. Based on the research in previous research papers, suitable consistent method is used to expand or shrink the neural network size until a reasonable output is obtained. In this work we tried different sizes for the neural network using python and we found that the best among them .. Different kinds of Neural Networks are used for various purposes, for example for predicting the sequence of words we use Recurrent Neural Networks more precisely an LSTM, similarly for image classification we use Convolution Neural Network. This algorithm will detect the face emotions and sends the mail to the user during abnormal face emotions were found.

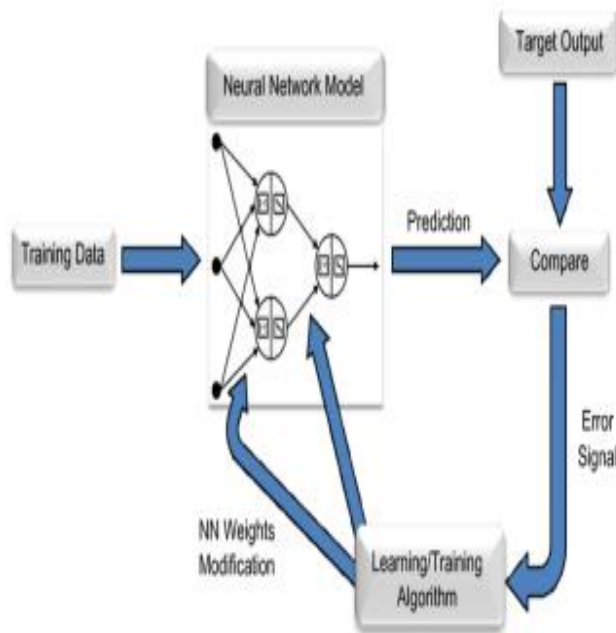


Fig .4.working of CNN.

6. Alert Message

The SMTP protocol has been accustomed to send the mail to the user. SMTP is an application layer protocol .The local network intends to send an E-mail message to another address using the local server, or when the E-mail messages are transferred through this protocol. The SMTP server is usually on listening mode. The procedure of SMTP client/server can be summarized in establishing of a TCP connection with the E-mail server on port 25, starting with the client/server commands dialogues procedure, transferring the E-mail message to the sender's E-mail sever and finished by quitting the communication with the E- mail server on port 25.The client/server commands dialogue and the E-mail message data will be sent to the E-mail server using the established connection.

For the SMTP protocol, the minimum E-mail server response time taken to accomplish the processes of creating TCP connection session, receive E-mail server message, send the command and terminate the connection was 10 seconds. After successfully establishing the TCP connection the client process sends the mail.

IV. RESULT AND ANALYSIS

Our approach ensures full automatic solution to identify human expressions as well as rising above facial expressions variation and intensity. The performance of the proposed approach was evaluate and differentiate the expression. Since we are interested to recognize seven facial expressions, we consider only the specific rules for each expression. The of facial expression classes is presented as: (1) JOY, (2) SMILE, (3) SADNESS, (4) SURPRISE, (5) ANGER, (6) DISGUST, (7) FEAR and (8) Unknown Expression. Experiments show encouraging results in recognizing seven prototypic facial expressions. The performance of our approach can be further improved by integrating other more precise active contour technique to extract accurately facial features shape. This being so, we intend to conduct further studies on more different facial expression databases.

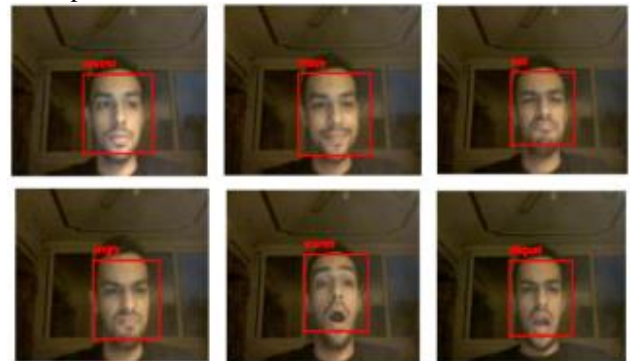


Fig.5. Results of proposed system.

V. CONCLUSION

This work presents the detection of abnormal facial expressions from the live video obtained. The abnormal facial expressions are detected by comparing the facial expressions obtained from video with the standard dataset collected from. This is done by using convolutional neural networks. This work is initiative to help the patients as well as people who are living alone. By using this project we can easily detect the abnormal facial expressions and report to the doctor for emergency needs. We use convolutional neural networks to classify the facial expressions and matching with the pre-defined data set to check the condition of the patient. The experiments show that the suggested algorithm has performed more accurate results in compared to the previous implemented algorithm.. The data set was already developed by data scientists with accuracy of 92.4%.To improve the accuracy of the project we implemented the CNN.

Efficiency and performance of the result is improved compared to the existing model.

Trans. Multimedia ,Oct. 2016 ,pp. 1-13 doi 10.1109/TMM.2016.2617741

REFERENCES

- [1]. M. Song et al., "A generic framework for efficient 2-d and 3-d facial expression analogy," IEEE Trans. Multimedia, vol. 9, no. 7, pp. 1384–1395, Nov. 2007.
- [2]. Akshay Asthana, Miles Delahunty, Abhinav Dhall, and Roland Goecke, "Facial performance transfer via deformable models and parametric correspondence," IEEE Trans. Vis. Computer Graphics, vol. 18, no. 9, pp. 1511–1519, Sep. 2012.
- [3]. K. Li, F. Xu, J. Wang, Q. Dai, and Y. Liu, "A data-driven approach for facial expression synthesis in video," in Proc. IEEE Conf. Computer Vis Pattern Recognition., 2012, pp. 57–64.
- [4]. K. Li et al., "A data-driven approach for facial expression retargeting in video," IEEE Trans. Multimedia, vol. 16, no. 2, pp. 299–310, Feb. 2014.
- [5]. L. Xiong, N. Zheng, Q. You, and J. Liu, "Facial expression sequence synthesis based on shape and texture fusion model," in Proc. IEEE Int. Conf. Image Process., Oct. 2007, vol. 4, pp. IV-473–IV-476.
- [6]. S. Agarwal and D. P. Mukherjee, "Decoding mixed emotions from expression map of face images," in Proc. Int. Conf. Workshops Automatic Face Gesture Recognition., Apr. 2013, pp. 1–6, doi. 10.1109/FG.2013.6553731.
- [7]. Hector P. Martinez, Georgios N. Yannakakis, John Hallam, "Don't Classify Ratings of Affect; Rank Them!" IEEE Trans. Multimedia, vol. 5, no. 3, pp. 314–326, Sep. 2014.
- [8]. Yuwen Wu ,Hong Liu ,Hongbin Zha, " Modeling Facial Expression Space for Recognition", supported by National Natural Science Foundation of China (NSFC). Project No: 60175025
- [9]. Charles Malleson, Jean-Charles Bazin, Oliver Wang, Derek Bradley , Thabo Beeler , Adrian Hilton , Alexander Sorkine-Hornung, "Face Director: Continuous Control of Facial Performance in Video" , IEEE International Conference on Computer Vision, Nov. 2015 , pp. 3979-3987 , doi. 10.1109/ICCV.2015.453
- [10].Wei-feng Liu, Ji-li Lu, Zeng-fu Wang, Hua-jun Song , "An Expression Space Model for Facial Expression Analysis", IEEE Congress on Image and Signal Processing, Nov. 2008 , pp. 680-684 , doi 10.1109/CISP.2008.216
- [11].Sicheng Zhao, Hongxun Yao, Yue Gao, Rongrong Ji, Guiguang Ding," Continuous Probability Distribution Prediction of Image Emotions via Multi-Task Shared Sparse Regression", IEEE