International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 04 | Apr-2018 www.irjet.net p-ISSN: 2395-0072

Automating the Human Resource System - A modular ICSA approach to Human Recruitment via Dynamic Emotional Evaluation.

Ajay G[1], Deepu R[2]

[1]*Mtech. Student, Department of Computer Science and Engineering, Maharaja Institute of Technology, Mysore, Karnataka, India*
[2]*Head of Department for Computer Science and Engineering, Maharaja Institute of Technology, Mysore, Karnataka, India*

Abstract - *A need exists to augment human potential to create an environment of trust and co-operation. Humans are known for their ability to lie and present biased facts. Thus the analysis of this important social cue of facial emotions is vital in order to identify and rectify any discrepancy in the behavior of the person in front of us. The use of Viola-Jones Algorithm, Haar Cascades, Eigen-faces algorithm also the use of parameter-based segment identification can be used to create a mechanism to detect and classify emotions. This is an ICSA (Intuitive Computer Sensed Analysis) Approach.*

***Key Words*: OpenCV, HAAR Cascades, ICSA TOOL, HR, Automation, Emotional Analysis, Computer Vision, AI**

1. INTRODUCTION

Facial Recognition is a major advancement in the field of computing for the year 2017 and beyond. As real time detection of human characteristic also embodies the desire of the human species to achieve a goal of producing a machine that one day can identify human faces and the emotions behind the faces. This will allow for reacting appropriately to situations, and support the decisions being made. The population will change in the coming time as the species may evolve over several millennia alongside our bodies. But stability is found in the presentation layer, the dimensionality of the face that is used to express the sentiment or bias. But that is not all, people also react to not just people but also to certain non-verbal cues such as a situation of being in the presence of those who they appreciate and care for and those they do not care for. The entirety of this project's premise is to create a mechanism or a chain of processes that convert a sentiment expressed into a emotion and match it t the right state and textually convey the outcome of a statistically major occurrence of emotion and classify the state of a person over a duration of time as a state of being. Most people possess a set of emotions commonly in inheritance to the surroundings and observations such as happiness, sadness, excitement, anger, distraction, attentiveness and many more. These can be considered states of being. Point of interest is to get an overview of sentiments expressed over a period of time.

The equation to state this change in expression over time is given by the following :

$$d(x)/d(y) = (\sum(X1+.....+Xn) / N) * (\theta c)$$

Where X is the point of sentiment, N is sentiment expression count, θc is the change factor it is a value based on the change rate of our face in unit time, d(x) change in expression, d(y) change in time-step.

1.1 Existing System

The existing system is based on manual human emotional analysis of a person by another person. There are some existing options but they all base their proposition on the use of cloud-based tools and third-party APIs which can be inefficient in a location without a good internet connection and also do compromise ownership of your data as the ability the control the storage of the data when it is passed server to server in order to get processed is lost to the end user. This is costly and time inefficiency and also a very mistake prone and at time very un-insightful approach. With a lot of variables that are not under human control, such as lack of attention from the interviewer to detect the minute changes in the face of the person being interviewed, improper identification of the person being interviewed, wrong documentation. All of these can be potential business risks and future failures if not dealt with in time.

1.2 Proposed System

The system is an automatic video based analysis engine whose development is predominantly done in Python and Tcl/Tk. It features a use of the eigenface, viola-jones, harr-cascade which are popular ML feature extraction and identification algorithms, in order to identify the facial features of the person in the picture frames derived from the time-step based division of the input video file in mpeg4 format.

2. RELATED WORK

Latifa Greche et.al [1], have proposed a system that is essentially based on histograms of oriented gradient features and multilayer feed forward neural network. Results demonstrate that good recognition rates can be obtained using small image resolution in both color spaces grayscale and RGB. The performance was evaluated on three different datasets by varying two parameters: face resolution and color space of images. Works using Viola-Jones Algorithm and HOG feature extraction. Weixuan Chen et.al [2], have proposed a system that takes into account vital signs, cognitive load, and stress can be remotely measured from human faces using video-capturing devices under ambient light, which raises both wide applications and privacy issues. To avoid the immoral use of this technology, there is a need for methods to eliminate physiological information from facial videos without affecting their visual appearance. A custom novel algorithm based on motion component magnification that inputs a video and outputs its replica with physiological signals removed. The paper is about eliminating non-required features to prevent misuse but it can be used to eliminate non-essential features from interfering in the accuracy of identification. Min Luo, Yadong Luo et.al [3], have proposed a system based on OpenCV in c++ it is a low light photo acquisition algorithm that captures 3D images in low light conditions based on super resolution. Considers and critiques the boosting algorithm for its complicated approach. Uses the 3D face recognition algorithm built on opencv using haar-cascade. Experiments show that the design scheme with a real-time focusing speed at 0.05 Seconds (single acquisition), the facial recognition rate of 99.3%, Proves that all algorithms don't work well for all conditions. There is a disadvantage that is not mentioned the space occupancy for the newer algorithm is huge compared to the previous algorithm.

Asma El Kissi Ghalleb et.al [4], have proposed a system that is based on fusing two soft modalities of the face, which are the skin and hair colors, Six soft modalities of the body – one for the height, four for the body measurements, and one for gait cycle. Eigenface, Viola-Jones, sparse rrepresentation and random subspace method and geometric approach are used to detect height, gait and face features. Vamshi Krishna Gudipati et.al [5], have proposed a system for facial expression recognition using adaboost, haar cascades on mouth detection and logistic regression, python with OpenCV. The term facial expression recognition implies the order of the facial elements in one of the six fundamental feelings: joy, misery, repugnance, trepidation, outrage and astonish. Accuracy is for 28x10 pixel images. Limited feature extraction and accuracy is volatile. The image set is small at 43 and 64 images respectively. Rajesh K M et.al [6],have proposed a system where the key elements of the face are considered for prediction of face emotions and the user. The variations in each facial feature are used to determine the different emotions of the face. Machine learning algorithms are used for recognition and classification of different classes of face emotions by training of a different set of images. The proposed algorithm is implemented using open-source computer vision (OpenCV) and machine learning with python. Fisher face algorithm presents a highly accurate approach for face recognition; it performs two classes of analysis to achieve recognition i.e. principal component analysis (PCA) and linear discriminant analysis (LDA).

Alessandra Bandrabur, Laura Florea and Corneliu Florea et.al [7], have proposed a system based on face muscles which can produce a number of 46 basic facial actions named Action Units (AU) and all the fundamental expressions are produced by combinations of AUs. The approach is extensively tested on the facto emotion database, namely Cohn-Kanade+. The AUs are dynamic and have three phases: onset, apex and offset. pre-processing, feature selection and extraction and classification. Facial recognition system uses Haar Cascade Classifier. Face fiducial points location-based approach. Loredana Stanciu et.al [8], have proposed a system based on haar cascade and FACs used along with the similar and known concept of AUs. The paper mentions the system as a concept with a layout, it is presented however in generalized terms the software does not exist and doesn't mention creating of any insight report or a set creator to address the problem of training data and acknowledges that the detection of emotion is a difficult task in a 2D space let alone a 3D one. The part where the author had difficulties was obtaining hundreds of pictures to help with the training with the same data object (person) and the training of the system in order to detect the emotional state. The application is used to detect emotional states based on facial features, specifically, according to face shape, position, and shape of the mouth and nose.

Weihai Chen et. al [9], have proposed a new algorithm that is supposed to synthesize such an effect for a single-portrait image. The foreground portrait is detected using a face prior based salient object detection algorithm. Then with an improved gradient domain guided image filter, the details in the foreground are enhanced while the background pixels are blurred. In this way, the background objects are defocused and thus the foreground objects are emphasized. The resultant image looks similar to an image captured using a camera with a large aperture. Face prior based saliency detection algorithm and a new gradient domain guided image filter (GGIF). The saliency detection algorithm in is designed for ordinary images. The GGIF can preserve edges better than both the guided image filter (GIF) in and the weighted GIF(WGIF). Mehryar Emambakhsh and Adrian Evans [10], have proposed a system that launches a investigation to go into the 3D nasal region for human identity authentication and verification purposes and presents a novel algorithm that provides very high discriminant strength, comparable with recent 3D face recognition algorithms, which use the whole facial domain.The algorithm first finds an approximate location of the nose tip and then finely tunes its location, while accurately determining the position of the nasal root and detecting the symmetry plane of the face. Next, the locations of three sets of landmarks are found: subnasale, eye corners, and nasal alar groove. These landmarks are utilized on feature maps created by applying multi-resolution Gabor wavelets to the surface normals of the depth map. Two types of feature descriptors are used: spherical patches and nasal curves. Feature selection is then performed using a heuristic genetic algorithm (HGA) and, finally, the expression-robust feature descriptors are applied to the well-known and widely used 3D Face Recognition datasets. An additional advantage of the proposed approach is that a fast Principal Component Analysis (PCA)-based self-dependent method can be employed for facial pose correction. This eliminates the need for sophisticated pose correction algorithms or reference faces for fine tuning the alignment. Michael Xuelin

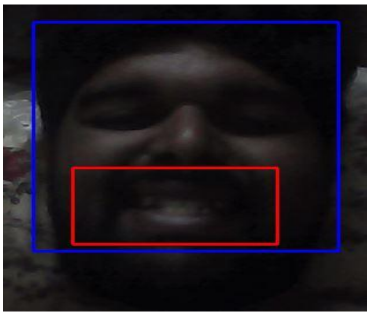
Huang Et al. [11], have proposed a system where the case of the PADMA (Personalized Affect Detection with Minimal Annotation), which is a user-dependent approach for identifying affective states from spontaneous facial expressions without the need for expert annotation is examined. The conventional approach relies on the use of keyframes in recorded affect sequences and requires an expert observer to identify and annotate the frames. It is susceptible to user variability and accommodating individual differences is difficult. An alternative is a user-dependent approach, but it would be prohibitively expensive to collect and annotate data for each user. PADMA uses a novel Association-based Multiple Instance Learning (AMIL) method, which learns a personal facial affect model through expression frequency analysis and does not need expert input or frame-based annotation. PADMA involves a training/calibration phase in which the user watches short video segments and reports the effect that best describes his/her overall feeling throughout the segment. The most indicative facial gestures are identified and extracted from the facial response video, and the association between gesture and affect labels is determined by the distribution of the gesture overall reported effects. Hence both the geometric deformation and distribution of key facial gestures are specially adapted for each user.

Jiajia Lei et. al [12], have proposed a multibiometric system using face and ear features are increasingly adopted for forensic and civilian applications to address the challenges of facial expressions and occlusions. Although numerous ear and face recognition techniques have been proposed, not much work has been conducted in the field of 3-D fiducial points localization and 3-D ear detection. Presently it is observed that an effective and efficient system of ear landmark localization, ear detection, and pose classification based on 3-D ears captured under large yaw variations is used. By utilizing the symmetrical property of human heads and classifying the ear with respect to its pose, all three tasks can be fulfilled given either left or right ears, without any prior pose information. A novel ear tree-structured graph (ETG) is proposed to represent the 3-D ear, after which a 3-D flexible mixture model is trained to locate the landmarks automatically. Then, the ear region is segmented based on them and the pose of the ear, i.e., whether it is a left or right ear, is classified based on the detected ETG. To the best of our knowledge, this paper is the first to present automatic landmark localization of 3-D ears extracted from facial scans with significant pose variations. Experiments were conducted at the University of Notre Dame collection F, G and J2, which contain large occlusion and pose variations, validating the effectiveness of the proposed methods. Zhenwei Miao et. al [13], have proposed a system based on the Laplacian of Gaussian (LoG) filter is widely used in interest point detection. However, low contrast image structures, though stable and significant, are often submerged by the high contrast ones in the response image of the LoG filter and hence are difficult to be detected. To solve this problem, we derive a Generalized LoG (GLoG) filter and propose a zero norm LoG filter. The response of the zero-norm LoG filter is proportional to the weighted number of bright/dark pixels in a local region, which enables this filter to be invariant to the image contrast. Based on the zero-norm LoG filter, we develop an interest point detector to extract local structures from images. Compared to the contrast dependent detectors, such as the popular Scale Invariant Feature Transform (SIFT) detector, the proposed detector is robust to the illumination change and abrupt variations of the image. Experiments on benchmark databases verify the superior performance of the proposed zero-norm LoG detector in terms of the repeatability and matching score of the detected points as well as the image recognition rate under different conditions.

Jacinto C. Nascimento [14], have proposed a new combination of deep belief networks and sparse manifold learning strategies for the 2D segmentation of non-rigid visual objects. With this novel combination, we aim to reduce the training and inference complexities while maintaining the accuracy of machine learning based non-rigid segmentation methodologies. Typical non-rigid object segmentation methodologies divide the problem into a rigid detection followed by a non-rigid segmentation, where the low dimensionality of the rigid detection allows for a robust training (i.e., a training that does not require a vast amount of annotated images to estimate robust appearance and shape models) and a fast search process during inference. Therefore, it is desirable that the dimensionality of this rigid transformation space is as small as possible in order to enhance the advantages brought by the aforementioned division of the problem. It is proposed to use of sparse manifolds to reduce the dimensionality of the rigid detection space. Furthermore, it is proposed that the use of deep belief networks to allow for a training process that can produce robust appearance models without the need of large annotated training sets. A test is performed on the approach in the segmentation of the left ventricle of the heart from ultrasound images and lips from frontal face images. The experiments show that the use of sparse manifolds and deep belief networks for the rigid detection stage leads to segmentation results that are as accurate as the current state of the art, but with lower search complexity and training processes.

They require a small amount of annotated training data. Nelly Pustelnik et. al [15], have proposed a system for texture segmentation constitutes a standard image processing task, crucial for many applications. The present contribution focuses on the particular subset of scale-free textures and its originality resides in the combination of three key ingredients: First, texture characterization relies on the concept of local regularity ; Second, estimation of local regularity is based on new multiscale quantities referred to as wavelet leaders ; Third, segmentation from local regularity faces a fundamental bias-variance trade-off: In nature, local regularity estimation shows high variability that impairs the detection of changes, while a posteriori smoothing of regularity estimates precludes from locating correctly changes. Instead, the present contribution proposes several variational problem formulations based on total variation and proximal resolutions that effectively circumvent this trade-off. Estimation and segmentation performance for the proposed procedures are quantified and compared to synthetic as well as on real-world textures

3. SYSTEM IMPLEMENTATION



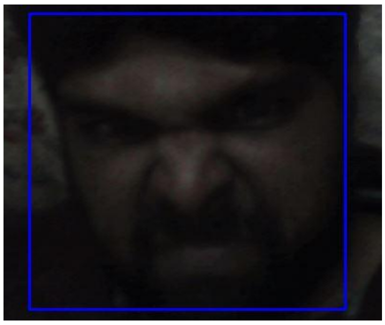
Picture - 1: Smile Detected (Positive)

Picture - 2: Smaller Variant of Smile Detected (Positive)

The red colored box around the lips has been drawn using the opencv cv2.rectangle function which allows us to draw shapes of rectangles around the article needed to be highlighted. In this case, it is the lips. Also, note that the box does not appear when the emotion detected is not a smiling face or not detecting the required emotional state of happiness.

Picture - 3: Angry Face not Detected as Happy Emotion (Negative Input)

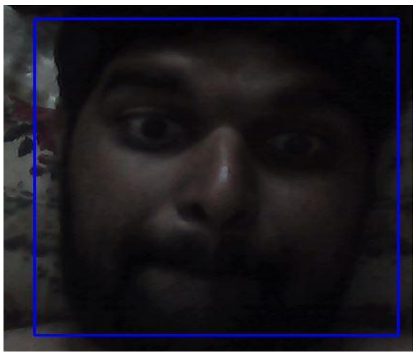
Do observe that the above examples clearly illustrate the versatility of the haar based novel classifiers. There is the option to train novel classifiers via the open architecture. The important thing to note however is that this process is not without its flaws. It does provide an area for improvement, it creates a lot of false positives when the person has glasses on his/her face when in a low light environment. The algorithm, however, has no issues performing well in low or well-lit conditions for the emotion of happiness based on the smile detection. Also, note that the emotions detected can be classified by varying the parameters to the respective detection value for the said emotion. For example, if a classifier focuses more on the eyes and its wide movements we can identify "shock" or if the parameters are said to focus on the nose of the subject we can gain insights to the emotion "anger". The feature most morphed during display of a emotion can act as an identification tool for the emotion.



Picture - 4: Angry Face (Suppressed) not Detected as Happy Emotion (Negative Input)

These examples clearly illustrate the working of the haar cascades that have been trained for this purpose of emotional state recognition in a low light condition.

Picture - 5: Shock Not Detected as Happy Emotion (Negative Input)



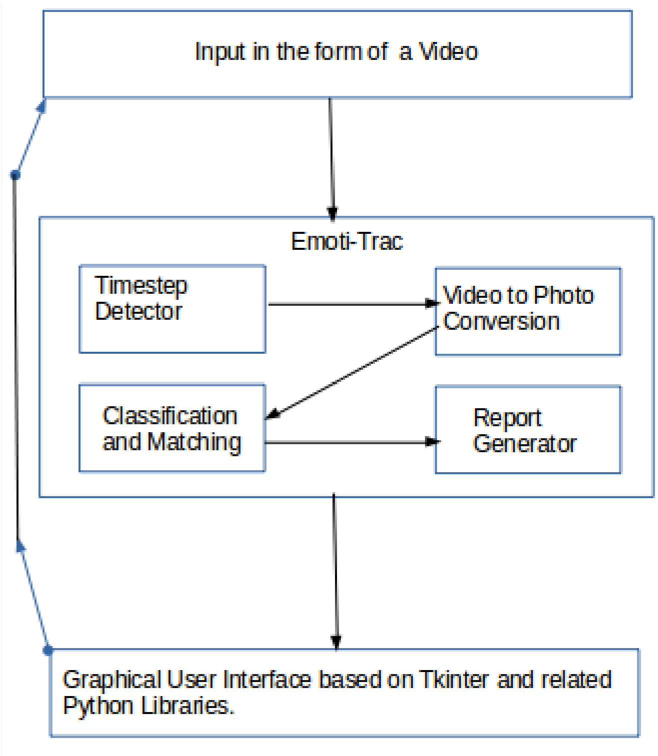
Picture - 6: Shock with Hidden Lips not classified as a Happy Emotion (Negative Input)

Picture - 7: Sadness Not classified as a Happy Emotion (Negative Input)

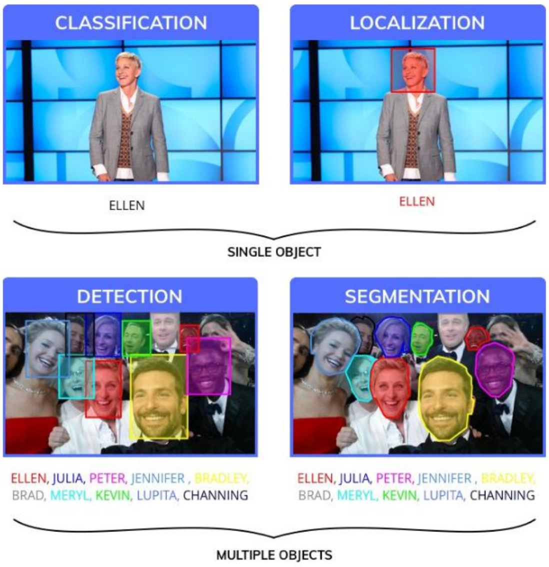
Also, do note that there are many alternative methods to achieve the same results as above, for example, one may use Bezier Curves or LBP classifiers present in the opencv library of pre-trained classifiers.Mainly to classify facial signatures and generate the emotional output as required.

The classifiers used in the above example were all trained separately and did not come pre-trained with opencv. The classifiers and method to train a novel classifier can be found on the www.github.com repos under "ICSA-HR" repo name. All work in this regard may be followed on the following link :

https://github.com/AjGumma/ICSA-HR.

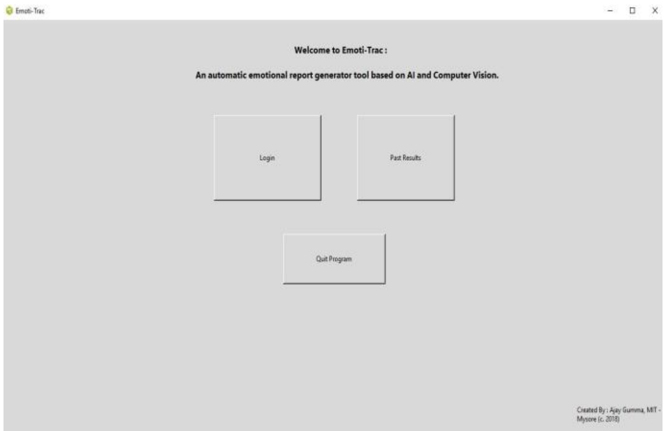


Picture - 8: System Architecture and flow



Picture - 9: Variations of Facial Identifications.

Picture - 9 depicts the differences in classification and localization and detection and segmentation. The pictures clearly have the distinct demarcation of the cv2 library based rectangle and segment functions.



Picture - 9: Tkinter Based GUI tool for the task of identifying emotions. (Main Menu)

3. CONCLUSION

The software tool eases two issues, one of improving the quality of people hired by better analyzing their emotional profiles and facial cues which help in making better decisions. It is a remote and easy to set up and use a system that gives results based on evidence-based data.

REFERENCES

[1] Histogram of Oriented Gradient and Multi-Layer Feed Forward Neural Network for facial expression identification, Latifa Greche, Najia ES-Sbai, Egons Lavendelis, IEEE 2017.