

A Survey: Software API and Database for Emotion Recognition

Renuka S. Deshmukh
Department of Computer Engineering
MIT
Pune, India
renukadeshmukh030@gmail.com

Prof. Vandana Jagtap
Department of Computer Engineering
MIT
Pune, India
vandana.jagtap@mit.edu.in

Abstract—Emotions are fundamental to human lives and their decision-making. Understanding an expression of emotional feeling between people forms an intricate web. There are systems, been developed that attempt to recognize aspects of emotion related behaviors and to respond to these, for example systems designed to improve the user experience or to change user behavior. The emotion recognition through facial images in recent times have proved to be an interesting topic for the researchers. Since two decades, a big research has been addressed to enhance Human Computer Interaction (HCI). There are a lot of applications and API-accessible software online that parallels the human ability to discern emotional behavior. The visual detection market is expanding enormously that can be seen noted from various system developed and described in this paper. This paper discuss various attributes, methods and emotional labels that are considered by various emotion API system. This paper also gives an overview of the databases available for inferring emotion through human facial features. The paper

Keywords—emotion; emotion recognition; image; facial emotion; database; facial emotion recognition tools.

I. INTRODUCTION

Emotions play a key role in day to day life when it comes to human interactions. Emotions are the form of nonverbal communication. Emotions are recognized by human through facial expression and are important clues for nonverbal communication among human beings. There are two major perspectives on the origin of emotions. According to one, emotions are said to be the products of natural selection. They are best understood using the explanatory tools of evolutionary psychology. In other words, emotions are socially constructed, and they vary across cultural boundaries. There is evidence supporting both perspectives [1].

Emotion recognition from facial images is an active research topic in human computer interaction (HCI). Facial Action Coding System (FACS) is a research tool useful for measuring any facial expression a human being can make [2]. It was developed in 1978, by Paul Ekman, along with Wallace Friesen and widely used even today [3]. The system is used to measure all visually distinguishable facial movements and for encoding, that results in the changes in the appearance of the face movements of facial muscles. It

is an anatomically based system for comprehensively describing all observable facial movement [2]. The observable component of facial movement is called an Action Unit or AU. FACS consists of 46 AUs of which 12 are for upper face, 18 are for lower face, and AUs 1 through 7 refer to brows, forehead or eyelids. The six basic emotions are anger, disgust, fear, happiness, sadness, and surprise [2].

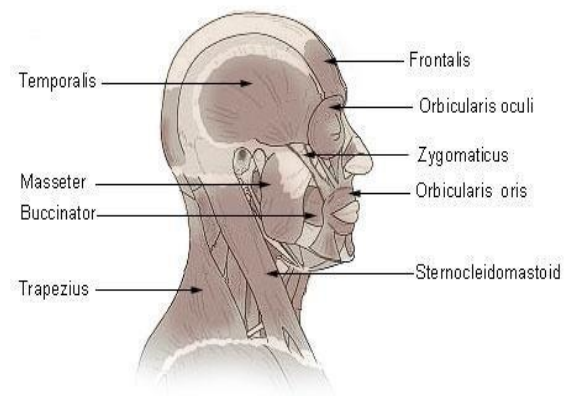


Fig. 1: Muscles of head and neck

Dr. Ekman developed the Facial Action Coding System (FACS) depicted below as a taxonomy of every human facial expression and the correlating muscles that form the expression.

The field of human computer interaction and affective computing focus on emotion recognition is a new frontier that could have huge consequences in not only advertising, but also in startups, healthcare, wearable's, education, and more. A lot of applications and API-accessible software online equals the human ability to discern emotive gestures. These algorithms driven APIs use facial detection and semantic analysis to interpret mood from photos, videos whether it be static or real time. Technology that reveals human feelings have can been implemented to spot struggling students in a classroom environment, or help autistics better interact with others. These applications may be helpful for the emerging researchers as a guideline to develop and take this area of research to a higher level of implementation. This paper

discusses various databases available and application to recognize human centric emotion.

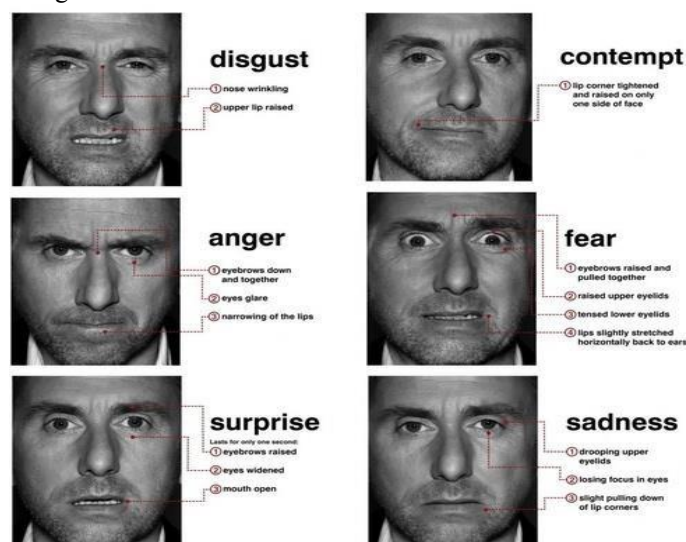


Fig. 2: Taxonomy of human facial expression

II. LITERATURE SURVEY

Emotion analytics is a very interesting blend of psychology and technology. The facial expression detection tools lump human emotion that can be categorized as Joy, Sadness, Anger, Fear, Surprise, Contempt, and Disgust. With this facial emotion detection, algorithms detect faces within a photo or video. It also senses micro expressions by analyzing the relationship between points on the face, based on curated databases compiled in academic environments. There are various tools developed for emotion detection and recognition. The Emotion API is a developer tool used to analyze face, image to determine the emotional reaction to the stimuli.

A. Microsoft cognitive services:

The tool includes the emotion API. The working starts by taking an input as a facial expression in an image, and eventually returns the confidence across a set of emotions for each face in the image. It also returns the bounding box for the face, using the Face API. If a user has already called the Face API, the user as an optional input can submit the

face rectangle. The emotions detected are anger, contempt, disgust, fear, happiness, neutral, sadness, and surprise. Computer Vision API describes images only in English, Face API detects up to 64 human faces in one image; these are some of the limitations. The software also provide additional API those are computer vision API; distill actionable information from images Face API, that Detect, identify, analyze, organize, and tag faces in photos. There are various API such as for face, video, emotion. Emotion API for Video provides aggregate results for the emotions of faces in a particular frame. The process of this API, first calculates emotion scores for each face in a video; done for smoothing the results over time for higher accuracy. It returns two types of aggregates result that include window mean scores; which gives a mean score for all the faces detected in a frame for each emotion. Then is the window face distribution, which gives the distribution of faces with each emotion as the dominant emotion for that face. Dominant emotions for each face have been determined based on the emotion with the highest score for that face. This is how the emotion in vedio is calculated through these steps. The key features of this API is recognizing emotion through facial image and in vedio.

These are some of the constraints that are needed the supported input image formats includes JPEG, PNG, GIF, BMP, Image file size should be no larger than 4MB, The detectable face size range is 36x36 to 4096x4096 pixels. Some faces may not be detected due to technical challenges, such as very large face angles (head-pose), large occlusion. Frontal and near-frontal faces have the best results Faces out of this range 36x36 to 4096x4096 pixels will not be detected. Reliability of the API decreases with partial frontal faces and Emotion API may not recognize emotions in images where the face is rotated 45 degrees or more. The API plans are as follows free plan for images 30K calls per month Video (uploads): 300 per month Video (query status): 3000 per month, Basic (Emotion API with Rectangles calls) plan for 10TPS images \$0.10 per 1000 calls, the Standard plan costs \$0.25 per 1000 calls Free for video, Images 10 TPS Video (uploads): 3000 per month Video (query status): 30K per month [4].

B. Kairos:

Kairos is a computer vision platform that helps companies radically change how they understand and interact with people. The algorithms, measure human faces in any camera-enabled app. Kairos Human Analytics is used by developers to transform their products that respond and react to people [5]. Kairos developed an Emotion API tool to analyze video to determine the emotional reaction to the stimuli found within the video. This Emotion API uses face detection and not face recognition. It can determine the emotion and attention of a person looking at the camera, but it does not attempt to identify who that person is in real life. It collects anonymous demographic and emotion data about the viewer(s). It works in recognizing emotion in real time. The key features are attention measurement, sentiment detection whether positive or negative, emotion recognition, gender detection and age detection.

Emotion API can detect facial expressions within the fraction of a second. The Kairos Emotion API can detect multiple faces up to 25 ft (8 meters) in an even lighting condition with a minimal powerful CPU. The faster the processor, the better the performance for collecting data at further distances. The Kairos Emotion API can detect emotion that includes smile, dislike, surprise, and attention. The demo version is free for limited access and further costs \$3000 per month and can be customized for unlimited access [5].

C. Eyeris EmoVu:

EmoVu created by Eyeris, is a facial detection product that incorporates machine learning and micro expression detection that allow an agency to accurately measure their content's emotional engagement and effectiveness on their target audience. The features include universal emotion detection, engagement metrics, mood indicator, age group, gender and a dashboard [6]. They do not rely on licensing third party components and have invested heavily into collecting their own proprietary datasets and own 100% of Intellectual Property. They have created novel techniques that use Deep Learning architectures, which employ Convolutional Neural Networks (CNN) to train algorithm and deploy it to the real world for continuous learning [7]. The application encompasses the emotion recognition results and this emotion recognition module produces an intensity measure for seven universal expressions. All expressions are normalized and sum to 1. The processing time is in milliseconds that is an integer value. It offers many advantageous factors such as intensity of negative and positive emotions, also the valence that can be described as the average of the maximum negative emotion intensity (includes Anger, Sadness, Fear, and Disgust) and the maximum positive emotion intensity (includes Joy and Surprise) that ranges from 0 to 0.5. The license key is required to get access. EmoVu offers wide platform support, including many tracking features, like head position, tilt, eye tracking, eye open/close, and more. The fig. 3 depicts the framework, where initially the face is captured and

further passed to Convolutional Neural Networks (CNN) for further classification of emotions [6].

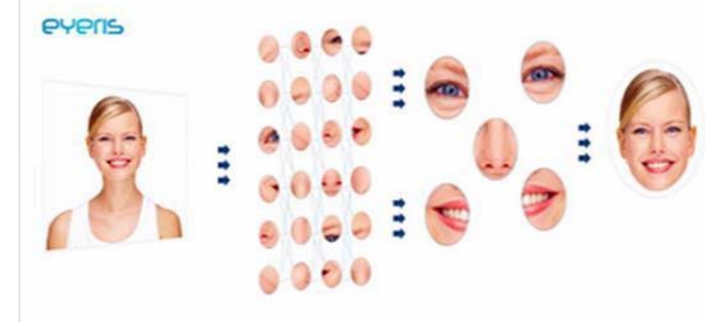


Fig. 3: Facial Feature Recognition

D. Sky biometric application:

If we wish to integrate face detection, face and attributes detection such as meta information like gender, emotions, face and facial features detection (nose, mouth, eyes). The Sky Biometry's API have made this easier than ever. Sky Biometry offers extremely saleable cloud based on-demand face recognition, detection and grouping. The algorithm based on deep neural networks, provides five times higher accuracy to identify full frontal faces and 10 to 15 times higher accuracy for unconstrained facial recognition. The improvements to the face recognition algorithm results in much higher accuracy in facial identification based on False Rejection Rate (FRR) same as the False Acceptance Rate (FAR) value. This improved users experience by resulting in fewer errors, and made the product easier to use and apply to a much broader range of face recognition applications, such as conducting automated facial image searches in large databases without the need for manual review. It includes emotions happy, sad, angry, surprised, disgusted, scared, neutral. It also returns attributes such as glasses and dark glasses by returning the value true or false. They are returned in the form of value, if the value is determined, it is returned along with confidence value as a percentage from 0 to 100. The subscriptions is a free, which is assigned with a limit of 100 face detection/recognition method calls an hour, monthly usage is limited to 5000 face detection/recognition method calls [8].

The tool results in faster face detection and more accurate estimation of facial attributes, including gender, smile, closed eyes, open mouth, glasses and dark glasses are also included [9].

E. Affectiva:

Affectiva developed Affdex technology that analyzes facial expressions to identify consumer's emotions such as whether a person is engaged, amused, surprised or confused. It employs advanced computer vision and machine-learning algorithms within a scalable cloud based infrastructure for identification of the emotions portrayed in a face video.

Affectiva has also developed SDKs for developing facial emotional analysis applications on both iPhone and Android device. Through this software development kit, developers can enrich digital experiences,

apps, games, and other technologies by integrating the emotion detecting software and allowing devices to respond to users emotions in real time without cloud support. In addition, they also provide a Restful API for apps that use cloud to analyze images and videos. It uses standard webcams, like those embedded in laptops, tablets, and mobile phones, in order to capture facial videos of people as they view the desired content [10].

Affectiva provides Emotion, Expression and Emoji metrics scores that indicates, when users shows a specific emotion or expression along with the degree of confidence. It considers total 0 to 33 indices for the facial landmark to detect emotions. The metrics can be thought of as detectors: as the emotion or facial expression occurs and intensifies, the score rises from 0 (no expression) to 100 (expression fully present). In addition, they also interpret a composite emotional metric called valence that gives feedback on the overall experience. Valence value ranges from 0 to 100 indicate a neutral to positive experience, while values from -100 to 0 indicate a negative to neutral experience. The emotions achieve accuracy in the high 90%. It is found that the classification performance is good as long as the subject's face occupies at least 30x30 pixels within an image. It recommends RGB camera and Minimum resolution 320x240. Also 10 frames per second minimum frame rate. Ideally the larger the face the better and assumes that the whole face includes eyebrows, lips and eye corners is within view of the camera. It generates six facial expressions Anger, Sadness, Disgust, Joy, Surprise, Fear and Contempt and also contains the world's largest repository of emotion data [11].

Affective includes Little Dragon, is a learning app made by Artha Ltd., a team of Psychologists, Educational Professionals and Game Designers that was found in 2015 Hong Kong. Little Dragon mission, made learning more enjoyable and more effective, by providing an educational tool that is both universal and personalized. Little Dragon uses the Affectiva SDK to detect facial expressions, body movements and position relative to the screen. The app analyzes this stream of data alongside other

signals to infer emotions that are relevant to the process of learning. Little Dragon's AI then adjusts to and anticipates the user's state to keep kids engaged. Field trials in rural China confirmed what the Artha team expected – kids like to be seen and have their achievements recognized. As Artha continues to refine Little Dragon's ability to follow and respond to the learners' emotional states they expect more engagement. Lessons are continuously adjusted to be at the right level of challenge. They are also excited by the ways in which they can integrate Little Dragon into existing classrooms and provide feedback to parents and teachers [11].

F. Emotient:

Emotient's co-founders have spent the past two decades innovating an automated emotion measurement technology based on facial expression analysis. The software is capable to work as a web-based service with any video camera or camera-enabled device, which could be a webcam, camera, tablet, or smart phone. The tool measures emotional responses via facial expression analysis. The approach used in Emotient combines proprietary machine learning algorithms, a self-optimizing data collection engine, and the state-of-the-art facial behaviour analysis to detect 7 primary emotions that includes joy, surprise, sadness, anger, disgust, contempt, and fear, as well as more advanced emotional states including confusion and frustration. It extract 19 Action Units directly and also extract data streams for confusion & frustration advanced emotions which is one of the advantage. The applications area include Consumer neuroscience and neuromarketing, Media testing & advertisement, Psychological research, Clinical psychology and psychotherapy, Medical applications & plastic surgery, Software UI & website design.

The system is also capable to detects all faces within the field of view of each frame and analyzes the facial expression. Emotient works with iMotions, which builds a full Attention Tool platform, which includes other biometrics features that includes eye tracking, heart rate measurement, and galvanic skin response (GSR) for improved academic, market, and usability research [11].

TABLE I. EMOTION RECOGNITION API TOOLS

Tool	Class label	Technique	Attributes	Number of face	Accuracy
Microsoft cognitive services	anger, contempt, disgust, fear, happiness, neutral, sadness, and surprise	Cloud-Based Emotion Recognition Algorithm.	Face rectangle top, left, Height, and width.	Individual face, Multiple faces	High
Kairos	smile, dislike, surprise, and attention	-	Eyes, eyebrows, nose, mouth		Moderate
Eyris EmoVu	facial detection	Deep Learning Architecture		Individual face, static image	-
Sky biometric application	facial attributes including gender, smile, closed eyes, open mouth, glasses and dark glasses	Deep Neural Networks	Closed eyes, open mouth, glasses and dark glasses.		-
Affectiva	engaged, amused, surprised or confused.	Machine-Learning Algorithms	Eyes, nose, mouth, eyebrows,	Individual face, Multiple faces, Real time images	90%
Emotient	joy, surprise, sadness, anger, disgust, contempt, fear,	Machine-Learning Algorithms	Eyes, eyebrows, nose, mouth		

	confusion and frustration				
OpenFace [12]	facial landmark detection and tracking: head pose detection, eye gaze estimation	Conditional Local Neural Fields (CLNF)	Face rectangle		95%
FaceReader 6.1 [13]	happy, sad, angry, surprised, scared, disgusted, contempt and neutral	Artificial Neural Network	Complete facial points.	Individual face, Real time images	95.1%
Nviso [14]	happiness, surprise, fear, anger, disgust, and sadness	Machine-Learning Algorithms	3D facial image feature	3D facial imaging technology	Highly accurate
Vision API [15]	joy, sorrow, and anger	-	Key facial features	Multiple face	Average

There are various emotion recognition API developed, that could be applicable in various fields of research applications. The most of the listed emotion API's have considered the six basic facial expressions, which are defined by Paul Ekman. The facial feature for emotion recognition include eyes, eyebrows, nose, mouth that are basic and are considered in most of the API's. The FaceReader and Nviso are some of the emotion API tools that consider overall facial points that result in higher accuracy. There are various approaches of classifying emotions that include hybrid approach, soft computing approach and machine learning approach. The emotion API listed here have implemented machine learning approach that prove to be more effective for the purpose of classification. It can be stated that the deep learning approach that is implemented for feature extraction in the Eyeris EmoVu, allows fast learning, speed and accuracy for emotion recognition systems.

There are enormous numbers of database available for facial expression that would be helpful for the researchers in this field. The databases available are illustrated in the table format.

TABLE II. FACIAL EMOTION DATASETS

Database	Class	Type	Number of samples	Image resolution
CMU Facial Expression Database (Cohn-Kanade) [16]	Joy, Surprise, Anger, Fear, Disgust, And Sadness	Posed	200 subjects	-
Extended Cohn-Kanade Dataset (CK+) [17]	Neutral, Sadness, Surprise, Happiness, Fear, Anger, And Disgust	Posed; spontaneous smiles	593 image sequences (327 sequences having discrete emotion labels)	640* 490
Japanese Female Facial Expressions (JAFFE) [18]	Neutral, Sadness, Surprise, Happiness, Fear, Anger, And Disgust	Posed	213 static images	256* 256
FERG (Facial Expression Research Group Database)-DB [19]	Angry, Disgust, Fear, Joy, Neutral, Sad, Surprise	Frontal pose	55767	768x768
Radboud Faces Database (RaFD) [20]	Neutral, Sadness, Contempt, Surprise, Happiness, Fear, Anger, And Disgust	Posed	Three different gaze directions and five camera angles (8*67*3*5=8040 images)	681*1024
Beihang University facial expression database [21]	Despair, Grief, Worry, Surprise, Flurry, Horror, Disgust, Fury, Fear, Doubt, Impatience, Hate, Contempt, Disparagement, Sneer, Smile, Plea, Laugh	-	-	-
NVIE (Natural Visible and Infrared facial Expression)[22]	Basic Expressions	Posed & Spontaneous expressions	-	-
The Yale Face Database [23]	Sad, Sleepy, Surprised	Posed	15 subjects	
The Psychological Image Collection at Stirling [24]	Smile, Surprise, Disgust	Posed	Aberdeen: 116 subjects Nottingham scans: 100 Nott-faces-original: 100 Stirling faces:36	-

III. CONCLUSION AND FUTURE SCOPE

Facial emotion recognition systems are paving ways in development of various applications that may be helpful for society. The objective of this research paper is to give a brief introduction towards various tools and description in terms of their features, the approach or methods through which they are implemented. A dataset provides a new and useful corpus, that are accompanied by the characteristics, for the benefit of the researcher in the field of human computer interaction. The paper also describes about the most commonly used facial databases that are considered in emotion recognition systems. The spontaneous expressions in some of the datasets, are targeted by the research community in order to develop a system that could adapt to the real time environment,.

To our knowledge there isn't any system yet developed that can capture the daily activities of the human and can be considered are the expansion feature in such systems. The study included in the paper would be most helpful to the researcher those are interested in the development of emotion recognition API system.

REFERENCES

- [1] J. Prinz, "Which Emotions Are Basic?", Emotion, Evolution, and Rationality, Oxford University Press, (2004).
- [2] Manual, F., Paul Ekman Group. Paul Ekman Group, LLC, (2016).
- [3] H. A. Alabbasi, F. Moldoveanu, A. Moldoveanu, "Real Time Facial Emotion Recognition using Kinect V2 Sensor", IOSR Journal of Computer Engineering (IOSR-JCE), Volume 17, Issue 3, Ver. II, 61-68 (May – Jun. 2015).
- [4] Microsoft.com. (2016), Microsoft Cognitive Services - Emotion API. Available at: <https://www.microsoft.com/cognitive-services/en-us/emotion-api>.
- [5] Kairos. (2016). Kairos API & SDK Documentation. Available at: <https://www.kairos.com/docs/>
- [6] B. Doerrfeld, 20+ Emotion Recognition APIs That Will Leave You Impressed, and Concerned | Nordic APIs |. Nordic APIs. Available at: <http://nordicapis.com/20-emotion-recognition-apis-that-will-leave-you-impressed-and-concerned/>, (2015).
- [7] EmoVu, EmoVu emotion recognition technology. Available at: <http://emovu.com/e/technology/>, (2016).
- [8] Available at: <https://skybiometry.com/documentation/>
- [9] SkyBiometry - Cloud-based Face Detection and Recognition API. (2016). New version of algorithm release - SkyBiometry - Cloud-based Face Detection and Recognition API. Available at: <https://skybiometry.com/new-version-algorithm-release/>
- [10] Lawton, G., Emerging Technologies Promise to Quantify Emotions. [online] Torque. Available at: <http://torquemag.io/2014/09/emerging-technologies-promise-quantify-emotions/> (2014).
- [11] Affectiva, Little Dragon - Affectiva. Available at: <http://www.affectiva.com/company/case-studies/little-dragon/> and <http://developer.affectiva.com/> (2016).
- [12] T. Baltrusaitis, P. Robinson, L. Morency, OpenFace: an open source facial behavior analysis tool, Applications of Computer Vision (WACV), IEEE, (2016).
- [13] L. Loijens, O. Krip, whitepaper "FaceReader Methodology", http://www.noldus.com/files/file_manager/downloads/whitepaper/FaceReader_Methodology.pdf
- [14] "Say It To My Face! Applying Facial Imaging To Understanding Consumer Emotional Response", AMSRS Conference 2011 White paper available at: file:///F:/white%20paper/nViso_AMSRS_2011_3D_Facial_Imaging_Best_Presentation_Award.pdf.
- [15] Vision API, available at: <https://cloud.google.com/vision/>
- [16] http://vasc.ri.cmu.edu/idb/html/face/facial_expression/index.html
- [17] Available: <http://www.consortium.ri.cmu.edu/ckagree/>
- [18] Available: <http://www.kasrl.org/jaffe.html>
- [19] Available: <http://grail.cs.washington.edu/projects/deepexpr/ferg-db.html>
- [20] O. Langner, R. Dotsch, G. Bijlstra, and D. H. J. Wigboldus, S.T. Hawk, A. V. Knippenberg, "Presentation and validation of the Radboud Faces Database", Cognition & Emotion, Psychology Press, 24(8), 1377—1388, (2010). <http://www.psypress.com/cogemotion>, DOI:10.1080/02699930903485076
- [21] Available: <http://ee.buaa.edu.cn/zykj/teachers/mx>.
- [22] Available: <http://www.equinoxsensors.com/products/HID.html>.
- [23] <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>
- [24] Available: (<http://pics.psych.stir.ac.uk/>)