



# Recognition of Human Face Emotions Detection Using Computer Vision Based Smart Images

Sanaullah Memon<sup>1\*</sup>, Noor Hassan Bhangwar<sup>1</sup>, Dr. Abdul Hanan Sheikh<sup>2</sup>, Najamuddin Budh<sup>3</sup>

<sup>1\*</sup> Department of Information Technology, Shaheed Benazir Bhutto University Shaheed Benazirabad, Campus Naushahro Feroze, Sindh Pakistan

<sup>1</sup> Department of Information Technology, Shaheed Benazir Bhutto University Shaheed Benazirabad, Campus Naushahro Feroze, Sindh Pakistan

<sup>2</sup> Department of Mathematics and Statistics, Institute of Business Management Karachi, Sindh Pakistan

<sup>3</sup> Department of English, Shaheed Benazir Bhutto University Shaheed Benazirabad, Campus Naushahro Feroze, Sindh Pakistan

## ABSTRACT

Humans share a collection of basic and essential emotions that are expressed by facial expressions that seem to be consistent. The automated identification of human emotion in images will be possible due to an algorithm that detects, extracts, and evaluates these facial expressions. The Face detector and recognizer application is a desktop application used to recognize human face emotions by using a computer vision based smart images. It consists of human face detection picture boxes and considers an image as an original image. It climaxes a face skin color, finds high impact area and identifies different emotions of the face from an image. The results of the image are depending upon the separation of Eyes & Lips movement of person. This is done by comparing face embedding vectors. It finds the smart photos focused on computer vision for successful identification of facial emotions in terms of various modes of speech such as smiling, shocking and weeping.

**Key words:** Computer Vision, Images, Recognition, Detection, Emotions, Human Face, Machine Learning

## 1. INTRODUCTION

Computer Vision is a field of man-made brain power that instructs computers to decipher and comprehend the visual world. It is concerned with the theory behind artificial systems that extract information from images. Utilizing computerized pictures from cameras and videos and profound learning models, machines can precisely recognize and arrange objects and after that respond to what they "see". All such change of information can accomplished for achieving some specific objective [1]. A computer vision intention is concerned with the design and development of algorithms that enables computers to improve their exhibition after some time dependent on information for example from databases. Learning indicates changes in a framework that empowers a framework to do a similar task all the more productively

whenever. Machine learning is malleable in both image processing and computer vision yet it is to be more observable in computer vision. It furnishes systems with the capacity to automatically take in and improve as a matter of fact without being explicitly customized. A delicate prologue to computer vision is shown in Figure 1. Over last decades, human face location has been investigated generally due to the recent advances of its applications, for example, security access control, data recovery in numerous unstructured media database, and progressed Human Computer Interaction (HCI). The input pictures can be caught by means of a few gadgets, for example, cameras and they can be controlled by different Computer vision strategies. Face detection is one of the most significant steps in many picture preparing applications, particularly in face recognition because of they need to find face first to perceive and condense data about the given casing progressively applications [2]. Face recognition strategies can be classified as highlight-based, layout-based, or appearance-based. Highlight based strategies try to look through the territories of particular picture features, for instance, the eyes, nose, and mouth, and after that check whether these features are in a possible geometrical game-plan [3]. Layout based technique, for example, active appearance models (AAMs), can manage a wide scope of posture and articulation changeability. Normally, they require great instatement close to a genuine face and are thusly not appropriate as quick face finders. Appearance based techniques examine over small overlapping rectangular patches of the image scanning for likely face up-and-comers, which would then be able to be refined utilizing a course of progressively costly however particular recognition algorithms [4].

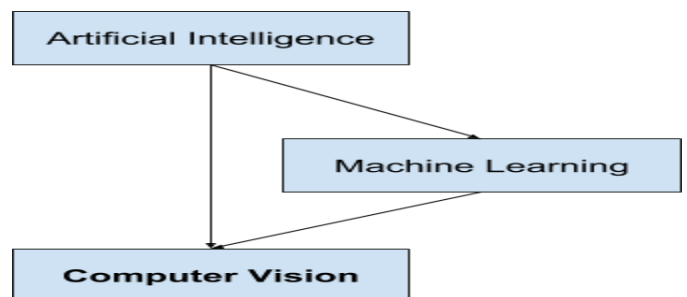


Figure 1: A gentle introduction to computer vision

## 2. LITERATURE REVIEW

Computer vision is a field that highlights systems to process, break down and get pictures. It expects to copy the capability of human vision by capturing a picture. One of the significant prevention to the uses of Computer vision is movement estimation, shape estimation and investigation. Gesture based communication is a computer vision based unblemished multi-faceted language that draws in signs formed by hand minutes in amalgamation with outward appearances and stances. It maps regular method for correspondence to human signs and motions empowering hearing debilitated individuals to impart among them [5]. Brilliant photography passes on the assurance of value improvement and usefulness augmentation in making stylishly engaging pictures. The vast majority of the present arrangements utilize a post handling methodology to decorate an image; the created device empowers a novel capacity of prescribing a decent look before the photograph is caught. Given a data face picture, the instrument normally evaluates the stance based in vogue score, finds the most appealing purpose of the face, and proposes how the stance should be adjusted [6]. Semantic highlights depict inborn attributes of exercises. Accordingly, semantics make the acknowledgment task increasingly solid particularly when similar activities appear to be outwardly unique because of the assortment of activity executions. A semantic space includes the most famous semantic highlights of a movement to be specific the human body (present and pose let), characteristics, related articles, and scene setting. Techniques are to be presented by exploiting these semantic highlights to perceive exercises from still pictures and video information just as four gatherings of exercises: nuclear activities, individuals' connections, human-object communications, and gathering exercises [7]. A head-pose recommendation system is presented that aides a client in how to best present while taking an image. Given an info face picture, the framework finds the most appealing edge of the face and proposes how the posture ought to be balanced. The proposal results are resolved adaptively to the appearance and beginning posture of the information face. The client study demonstrates the suggestion execution of the framework is tolerably identified with the level of congruity among the picture takers' proposals [8]. Self-Trackam analysis video edges caught continuously to confine human faces in each edge. The framework processes this data to situate the phone which is mounted on two servos equipped for panning on a level plane and tilting vertically. The region of interest is consequently situated in the focal point of casing utilizing controls got from faces positions [9].

A factual human shape model that depicts a body shape with shape parameters, a novel way to deal with consequently gauge these parameters from a solitary information shape outline utilizing semi- managed/supervised learning. By using outline includes that encode neighborhood and worldwide properties strong to clamor posture and view changes, and anticipating them to bring down dimensional spaces traversed multi-see learning with standard connection investigation [10].

Today machines are an important audience for any image. Rettberg looks at how facial acknowledgment algorithms examine our images for surveillance, confirmation of personality and better-altered business benefits, and relates this to understandings of machine vision as post-optical and non-authentic [11].

## 3. PROBLEM STATEMENT

Recognition of facial expression in the case of the real world is a long-standing problem which makes the process of extraction of features more complex. There are some major problems like face detection, face recognition, Image transformation and one of these problems is facial emotions detection that expresses the face expressions, identify pictures automatically during different expressions styles like smile, surprise, sad, cry and normal expressions by using computer vision based smart images.

## 4. AIM AND OBJECTIVES

The aim of the project is to investigate the facial expressions at the time of smiling, crying, surprising and normal by using computer vision technology based on smart images.

Objectives of the study are:

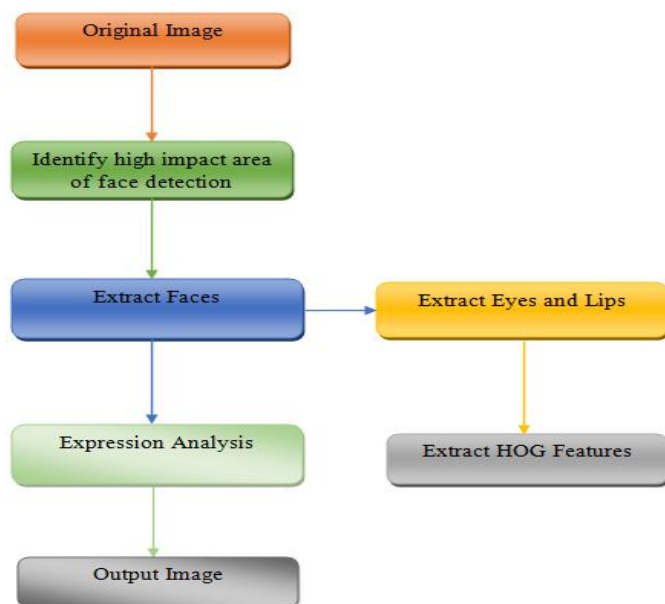
- Read lot of capabilities from the background study of Image detection and Image recognition.
- To study the image processing algorithms like facial feature recognition algorithm.
- To calculate the facial expression results during the project analysis.
- The reasons and factors behind the measured results are evaluated on the basis of the results of selected algorithms published.

## 5. RESEARCH METHODOLOGY

The Face detector and recognizer application is a desktop application and used for the human face detection & recognition. Research methodology is supposed to be experimental during application development phase. It takes an original image and identifies high impact area of face detection. It extracts facial features and emotions detections and analyses the expressions from an image. After that it compiles all the expressions in the desired output. This application is developed on visual studio with the use the Microsoft visual studio library OpenCV.

- Integrate OpenCV library to programming language like C#, JAVA or Python.
- Perform various operations like Face recognition, Facial expressions etc
- Implementation phase for image capturing by using facial expression such as smiling, weeping and crying.

- Do the system testing and release the deliverables in beta version.



**Figure 2:** Methodology for computer vision based smart images

A Viola-Jones cascade object face detector is used to remove the overall face from the image first. Using basic features known as Haar-like features, the Viola- Jones detection system attempts to recognize features of a face. Passing feature boxes is part of the procedure over an image and calculating the difference between adjacent regions' summed pixel values. The difference is then compared to a threshold, which determines whether or not an object has been identified. This necessitates thresholds that have been pre-trained for various feature boxes and features. The idea is that most faces and the features inside them can follow general requirements, so specific feature boxes for facial features are used. The Haar-like function method is extremely quick because it can measure the integral image of the image in a supervised manner and generate a summed area table. Then using just four values, the summed results of the pixels in any rectangle in the original image can be calculated. This enables several passes of various features to be completed quickly. A selection of features would be transferred for face recognition in order to detect certain aspects of a face, if one exists. The face is identified if enough thresholds are reached. The faces are removed and resized to a predetermined dimensional norm once they have been found. The average image for all of the training faces will then be determined. Faces that convey the basic emotions make up the entire training package. The mean image is then subtracted from the entire training collection of images. The scatter matrix  $S$  is then generated using the mean-subtracted training set. The goal is to find a change in basis that will enable us to express our face data in a more dimensionally optimized manner. As a result, the majority of the data can be retained as a linear combination of the much smaller dimension collection. PCA does this by attempting to minimize variation.

## 6. TOOLS AND TECHNOLOGIES

The Following tools and technologies are used during the development of project.

**C#:** It is general purpose programming language that incorporates strong scripting, declarative, procedural, abstract, object and component oriented disciplines. It was developed within the context of the .Net project by Microsoft headed by Anders Hejlsberg and his team and was accepted by ECMA and ISO.

**SQL Server:** Microsoft server is RDBMS that serves a wide range of enterprise IT database management, business intelligence and analytics applications. Microsoft SQL server along with oracle database and IBM's DB2 is one of the three leading database systems on the industry.

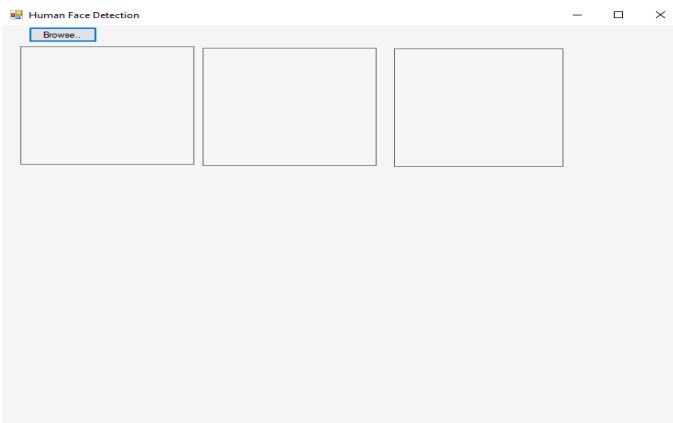
**OpenCV Library:** OpenCV is an open source computer vision and software library for machine learning. OpenCV was designed to provide a shared platform for the implementation of computer vision and to promote the use of machine perception in consumer products. Being a BSD-licensed software, OpenCV makes the use and alteration of the code simple for businesses.

## 7. EXPERIMENTS AND RESULTS

In this Facial Recognition project, the computer vision based smart images are captured of the person with different face expressions to identify the facial emotions. This is done by comparing face embedding vectors. The emotion Detection is classifying the emotions on the face of a person in distinct perspectives such as happy, angry, sad, normal, surprise, disgust or fear. AForge.NET is an open source C# platform developed for computer vision and artificial intelligence developers and researchers, neural networks, genetic algorithms, fuzzy logic, machine learning, robotics and so on. The framework consists of a collection of libraries and sample applications which display their characteristics:

- AForge.Imaging - library with image processing routines and filters;
- AForge.Vision - computer vision library;
- AForge.Video - set of libraries for video processing;
- AForge.Neuro - neural networks computation library;
- AForge.Genetic - evolution programming library;
- AForge.Fuzzy - fuzzy computations library;
- AForge.Robotics - library providing support of some robotics kits;
- AForge.MachineLearning - machine learning library;

The face recognition application is analyzed by the human face emotions. It consists of three human face detection picture boxes. In a first picture box, it considers an image as original image. In a second picture box, it highlights the face skin color and finds the high impact area of the image and in a third picture box, it searches the high impact face area and identify a face from an image. Following figure 3 shows the main screen of human face detection.



**Figure 3:** Human Face Detection Main Screen

Browsing plays an important role in computer vision and machine learning. An image browser is a piece of software designed specifically to enable viewing of digital image files. Modern digital photographers often have catalogs of tens of thousands of digital photographs, and an image browser can be used to seek, organize and delete images files quickly. Image browser allows the organization of images on the fly using the information stored in the metadata of the image file. It allows to extremely quickly find specific images or types of images, a much better solution to look through thousands of images one at a time. The main screen of the project shows the overall functionality of an image. First, browse the image in picture box. After browsing an image, the other picture boxes cover high impact area from an image and identify the human face. The result of the image is depending upon the Eyes & Lips movement of a person. It separates all the eyes and lips movement of person and show a result in terms of smile, cry, surprise, etc. Smile is often described as a smile emotional state characterized by feeling of happiness, pleasure, accomplishment and well-being. The following figure 4 shows the results of when a person is smiling.



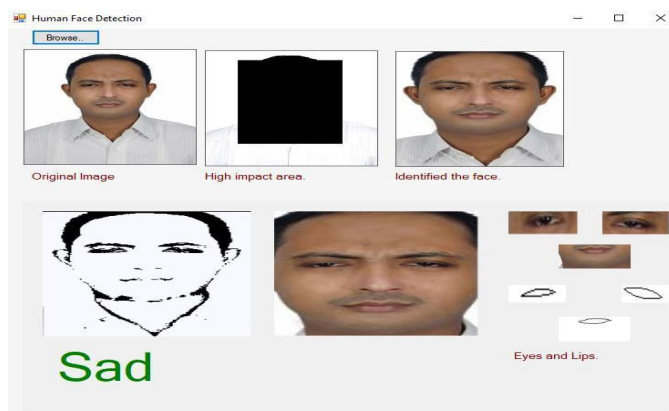
**Figure 4:** Smiling Human Face Expression

Surprise is one of seven universal emotions which show up when we hear sudden and unexpected sounds or gestures. His role as the briefest of the universal emotions is to concentrate our attention on deciding what is going on and whether or not it is risky. The following figure 5 shows the result of when a person is surprising.



**Figure 5:** Surprising Human Face Expression

Sadness is a form of emotion that is characterized by feelings of anger, hopelessness, disinterest and humidity. Sadness is universal emotion that everyone goes through at some point in their lives. In certain cases, people can experience persistent and severe sadness, which may lead to depression. The following figure 6 shows the result of when a person is sad.



**Figure 6:** Sad Human Face Expression

## 8. TESTING AND EVALUATION

**Primary data:** Primary data was obtained by firsthand (Primary source) information resulting in a collection of English-language questions related to the application.

**Secondary data:** Secondary data was also used for literature review, which justified our research work; throughout the light of various scholarly publications by various scholars.

**Sampling:** The data was collected using a random sampling method from 160 respondents, including both teachers and students of Shaheed Benazir Bhutto university Campus Naushahro Feroze.

**Statistical methods:** In starting, we checked the reliability of the instrument by using spss26 through Cronbach’s alpha.

### 8.1 VALIDITY/RELIABILITY ANALYSIS

The peer review and expert opinion offered validity for the research tool. After testing, several changes were made to the analysis instrument claims. To test the effectiveness of tools, the following final draft of the questionnaire was used as shown in Tables 1.

**Table 1:** Questionnaire for application

Qno: 1	Is this application reliable and easily accessible?
Qno: 2	Which feature do you think that it is missing and should be added to this project?
Qno: 3	Is this application providing accurate and optimized results?
Qno: 4	What do you think this application is efficient in marketplace?
Qno: 5	Does this application cover the needs that you want to acquire?
Qno: 6	Is this application said to be effective, efficient, and highly scoped in modern world technology?
Qno: 7	Does this application contain new idea and contribution from previous study?
Qno: 8	Are the factors behind the measured results are evaluated on the basis of the results of selected algorithms published?
Total Cronbach’s alpha: $\alpha = .78$	

### 9. CONCLUSION

Computer Vision provides the function to recognize an item as a particular object such as human being. In this study, it has been found that the face detector and recognizer application is totally worked on computer vision based smart images. It identifies the high impact area of the face that contains facial expressions in the images like happy, sad and surprise. It concludes that this application contains more effective approach for detection and recognition of the expressions from the images.

### REFERENCES

[1] Bradski, Gary, and Adrian Kaehler. Learning OpenCV: Computer vision with the OpenCV library. " O'Reilly Media, Inc.", 2008. <https://dl.acm.org/doi/book/10.5555/1461412>

[2] Sharifara, Ali. Enhanced Face Detection Framework Based on Skin Color and False Alarm Rejection. Diss.

Universiti Teknologi Malaysia, 2015. <http://eprints.utm.my/id/eprint/77598>

[3] Yang, Ming-Hsuan, David J. Kriegman, and Narendra Ahuja. "Detecting faces in images: A survey." IEEE Transactions on pattern analysis and machine intelligence 24.1(2002):34-58. <https://ieeexplore.ieee.org/document/982883>

[4] Szeliski, Richard. Computer vision: algorithms and applications. Springer Science & Business Media, 2010. <https://link.springer.com/book/10.1007/978-1-84882-935-0>

[5] Rao, G. Ananth, and P. V. V. Kishore. "Sign Language Recognition System Simulated for Video Captured with Smart Phone Front Camera." International Journal of Electrical & Computer Engineering (2088-8708) 6.5 (2016). <http://ijece.iaescore.com/index.php/IJECE/article/view/1144>

[6] Hu, Chuan-Shen, et al. "Virtual portraitist: An intelligent tool for taking well-posed selfies." ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 15.1s (2019): 1-17. <https://dl.acm.org/doi/10.1145/3288760>

[7] Ziaefard, M., & Bergevin, R. (2015). Semantic human activity recognition: A literature review. Pattern Recognition, 48(8), 2329-2345. <https://www.sciencedirect.com/science/article/abs/pii/S0031320315000953>

[8] Hsieh, Yi-Tsung, and Mei-Chen Yeh. "Head pose recommendation for taking good selfies." Proceedings of the Workshop on Multimodal Understanding of Social, Affective and Subjective Attributes. 2017. <https://dl.acm.org/doi/pdf/10.1145/3132515.3132518>

[9] binti Ariffin, Mariatul Kiptiah. "Self-Trackam." International Conference on Computational Intelligence in Information System. Springer, Cham, 2018. [https://link.springer.com/chapter/10.1007/978-3-030-03302-6\\_20](https://link.springer.com/chapter/10.1007/978-3-030-03302-6_20)

[10] Dibra, Endri, et al. "Shape from selfies: Human body shape estimation using cca regression forests." European conference on computer vision. Springer, Cham, 2016. [https://link.springer.com/chapter/10.1007/978-3-319-46493-0\\_6](https://link.springer.com/chapter/10.1007/978-3-319-46493-0_6)

[11] Rettberg, Jill Walker. "Biometric citizens: Adapting our selfies to machine vision." Selfie citizenship. Palgrave Macmillan, Cham, 2017. 89-96. [https://link.springer.com/chapter/10.1007/978-3-319-45270-8\\_10](https://link.springer.com/chapter/10.1007/978-3-319-45270-8_10)