

# Face recognition-based automatic hospital admission with SMS alerts

Mamata Parab · Ninad Mehendale\*

Received: date / Accepted: date

**Abstract** When the person met with an accident is brought to the hospital, there are many official formalities (e.g. admission form to be filled) before the treatment that can be started. In some severe cases, these formalities can delay the treatment, which could be fatal to the patient. The automated system which can fill these forms with the help of face recognition could severely cut down the delays. But, in some cases, the injury and the blood on to the face fail facial recognition. To overcome this problem, we have proposed a facial vector-based algorithm. In the current work, we have also demonstrated, sending the SMS to the concerned authorities (police) and even to the relatives of the patient automatically using GSM modules. The patient's information was received from centralized databases of different hospitals that are linked through the internet. We have tested the algorithm on more than 213k images from different databases like celebA, LFW, UCFI. We found that the maximum accuracy of our system was 98.23 %. As a proof-of-concept, we tried testing on 51 real-time patient images and found that the accuracy is 94.11 %. This automated form filling not only reduced the delay in hospital admission, but it also helped in treatment, because of the auto-filled medical history.

**Keywords** Face recognition · Hospital admission · Automatic form filling

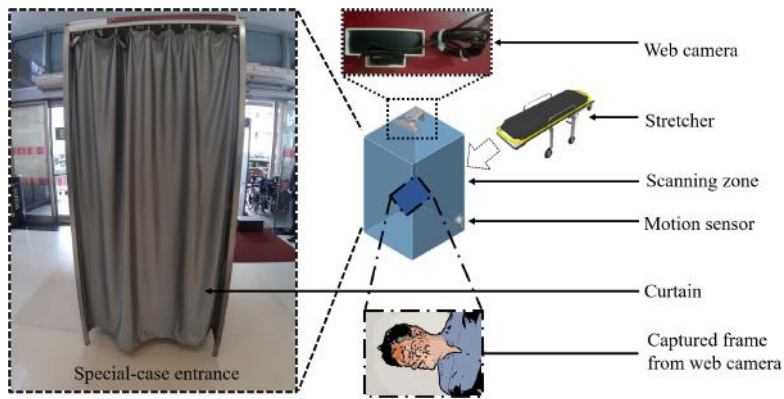
## 1 Introduction

When the patient has had an accident and brought to the hospital, it is critical to begin the treatment as soon as possible. [1]. Any delay in treatment may cause the

---

\* Corresponding Author  
N. Mehendale  
B-412, K. J. Somaiya College of Engineering, Vidyavihar, Mumbai, India  
Tel.: +91-9820805405  
E-mail: ninad@somaiya.edu  
ORCID id: <https://orcid.org/0000-0003-3037-5076>

M.Parab  
Ninad's research lab



**Fig. 1** System implementation diagram. A web camera is used to capture the video of the patient on the stretcher. We record this video only for patients entered through a special-case-entrance of the hospital. The special-case-entrance has a dark box-like structure, placed at the hospital entrance, that avoids the effects of external light. The motion sensor attached to the special-case-entrance allows time sync captured of the video when a stretcher entered the hospital.

patient's death. Informing the police is important to make sure that the accident case is genuine, and there is no conspiracy. According to a global status report on road safety of the year 2018 released by the World Health Organization (WHO), death due to road accidents has reached 1.35 million in a year [2]. That is almost 3699 people die because of the crash each day. Road accident is the eighth leading cause of death globally.

The severity of this road accident is much more in highly populated countries like India. According to the recent road accident statistics, almost 16 deaths happen per hour in India [3]. The high death rate in India is due to delay in starting treatment because of the various formalities that are mandatory before treating the accident patients. These formalities include admission form filling and sometimes informing the nearest police station. Informing to the police is important to make sure that the accident case is genuine, and there is no conspiracy.

The solution which may reduce the time taken for completing form filling formalities is the auto-filling of the form. This automatic form filling can be achieved using the identity recognition of the patient. Various studies involve the use of image processing in identity recognition. The most popular being biometric scans such as (fingerprint, iris, etc.) But generally, these scanners are costly. Furthermore, the patient's eyes are closed in most of these kinds of cases, making iris capture difficult. A simple, low-cost alternative to that is face recognition using simple webcam. Face recognition works on the basic principle that is based on facial features (skin color, moles on the face, etc.). Facial recognition becomes difficult if there is damage to the face during the accident.

Various types of accidents usually end up damaging the face (e.g. two-wheeler accidents without helmets, acid attacks, some domestic violence burnt cases). To overcome the problem in face recognition due to facial damage is to use face vectors. In general, the face vector is defined by different points that appear on the face after feature extraction in an ordered manner. For example, the simplest

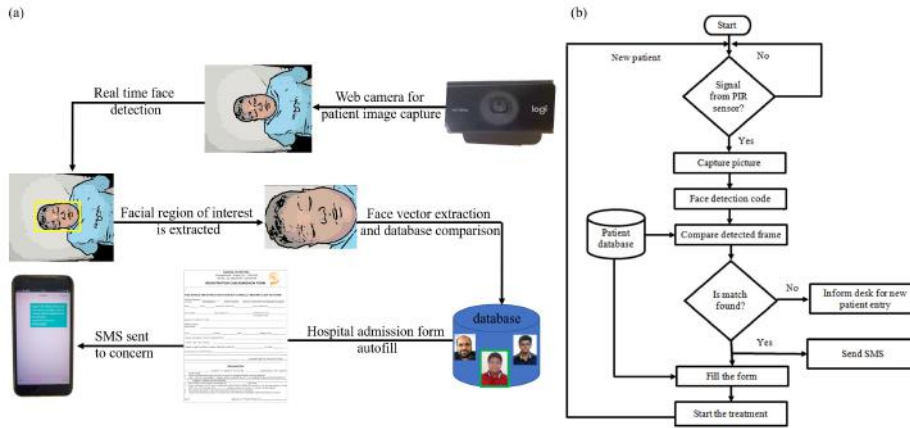
face vector could be the distance between two eyes, eyes and nose, eyes and lips, etc. in combination forms a unique set for each individual.

Unlike facial features, face vectors are rarely damaged during accidents. In cases when one or more facial vector is not available directly, it is still possible to reconstruct it because of facial symmetry. In the present work, we have developed a system that is capable of an automatic admission in the hospital using form filling with help of face recognition. The automatic form filling includes filling the data of the patient along with medical history (if available in the database). If the patient has a major injury or blood spot on the face, then also this proposed system can recognize the identity of the patient. The project has one more advantage of sending SMS to the nearest police station and relatives. As a demonstration of the proposed system has been tested with 51 patients. The system takes 1.5 min to scan the face of the patient, filling the form, and send the SMS. As there is a drastic reduction in time required in admission procedures the patient may get treatment early.

As shown in figure 1, a special-case entrance (i.e., dark box-like structure) was placed at the entry door of the hospital. The only unconscious states patients were passed through this special-case entrance. The special-case entrance was attached with a web camera that captured real-time video of the patient. The correct frame was extracted automatically from this captured video. The width of the special-case entrance was 1.2 meters, which confirmed the smooth passing of stretcher through it. The curtains were present at both ends of the special-case entrance to avoid interference due to outside light. The proposed system avoided the delay in completing hospital formalities in emergencies.

## 2 Literature review

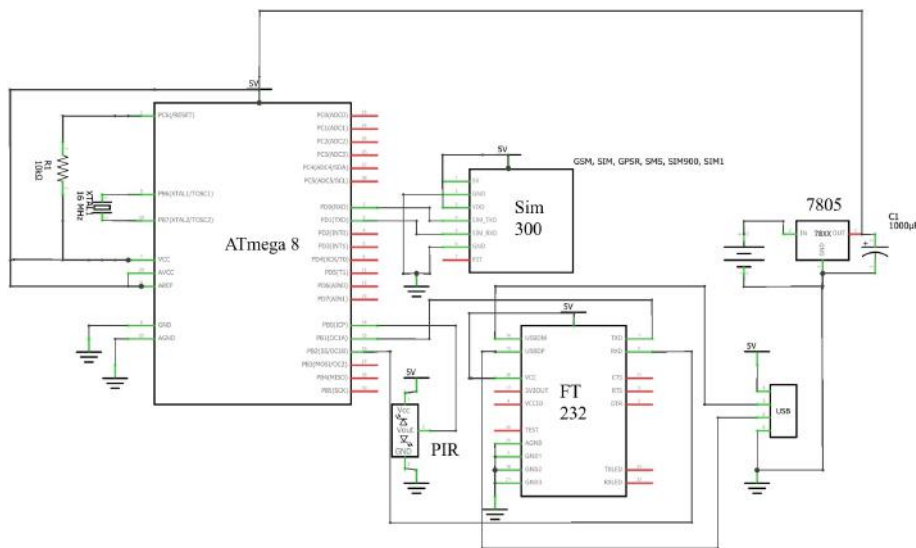
There are various studies done in the last few decades to improve face recognition algorithms. Recently, Luan Tran, *et al.* [4] published a work to minimize the large pose discrepancy. According to the authors of the paper, conventional techniques either perform face frontalization on an image or learn a pose-invariant representation face recognition. Singh *et al.* [5] conducted study on the facial injury due to road traffic accidents. They found that significant face damage caused to male face in contrast to the female face. Black *et al.* [6] compared injuries during sports and he found the effect of facial guards like a helmet on to the reduction of severity of these injuries. Canzi *et al.* [7] proposed facial recognition with trauma to optimize care and recovery after surgery. Trokielew *et al.* [8] performed dead body iris-based recognition and he was able to recognize people with 90% accuracy. Coulibaly *et al.* [9] provided survey on interpersonal violence face damage. Majumdar *et al.* [10] used VGG model to classify the domestic face violence from other same group people. He also extended their work for subclass contrastive loss for injured face recognition. [11]. McDonagh *et al.* [12] proposed a method for joint face detection by using cascaded regression of Hough-Transform. In the proposed technology, they scanned the image applied cascaded regression rather than passing the image through discriminatively trained filters. Yow *et al.* [13] proposed a feature-based algorithm for face detection. Feature selection was done based on the aspect ratio of the extracted object. Jung *et al.* [14] proposed a face detection method using local gradient patterns, by there study, they found that local gradi-



**Fig. 2** (a) System flow diagram. The video was acquired through a web camera. The face image is then extracted from the video. The face is detected and cropped from the image. The cropped face is then compared with the database image. If the match is found, then corresponding data filled into the admission form and SMS send to police and relatives. (b) Flow chart of the proposed system. The patient was passed through the black passage from the stretcher. Due to the motion of the stretcher, the PIR sensor was activated and send a signal to the web camera. Web camera record the video of moving stretcher. The desired image was extracted from the recorded video. The obtained image as input for the system and further enhanced by preprocessing. The processed image was used for face detection. The face feature vector was then compared with the database image. If a match is found then the hospital admission form filled automatically and SMS send to the concerned authority.

ent pattern for face detection is much better than local binary patterns. Mahajan *et al.* [15] proposed a face detection technique based on Histogram of Oriented Gradients (HOG) features. Their method used hierarchical knowledge-based model. The hierarchical knowledge-based method involved three levels for improving image quality. Images were filtered initially for easy background removal. The group of people proposed a novel technology for facial emotion recognition [16]. During implementation, we used the face vector technique in combination with convolutional neural networks. An enhanced face vector technique compared to literature is implemented in the proposed work. Parallely, the neural network has done an excellent job in the field of face detection and recognition. In most of the neural network approach, the input image has sub-sampled in different regions and converted it into a standard-sized sub-images. The sub-images are passed through the neural network filter. The filter improves the quality of an image by removing noise from it. Some of the work using the neural network technique has been done by Rowley *et al.* [17]. The performance of the system was very well for fronto-parallel faces, but results deteriorated when different views of the face were given as input. This technique failed for faces in different profiles. Labeled random graph matching approach in the cluttered image for detection of the face has been reported by Leung *et al.* [18]. The system included a technique for locating a quasi-frontal view of a face in a cluttered scene. In the system statistical model was used, which determined the distance between facial features. The statistical model was coupled with feature detectors. This technique was suitable for multiple views, but, at the same time, the technique failed to work under different imaging conditions. It was

difficult to robustly detect the facial feature for every subject because the structures of various facial features vary. Detecting the face in a complex background is a tedious task. Yang *et al.* [19] designed system to identify human faces in a complex background. This system used a hierarchical knowledge-based method. The hierarchical knowledge-based method involved three levels for improving image quality by filtering the background. The first two levels were based on mosaic images at various resolutions, and the third level was based on edge detection. This system gave excellent results for a fronto-parallel view of the face but failed for face detection in different profiles. Recently Deshpande *et al.* used the Viola-Jones algorithm and Fusion of PCA and ANN for face detection and recognition [20]. The system involved a method for recognizing the human face based on various features derived from the captured image. Their proposed method involved two stages. The first stage involved the detection of a face in an image using the Viola-Jones algorithm. The second stage involved the reorganization of the face, based on the fusion of Feed Forward Neural Network and Principle Component Analysis. The system provided better accuracy in recognition of a face. Similar to face detection a lot of recent work was also carried in automatic form filling. In 2010, Kadry *et al.* developed a system for wireless attendance marking systems based on iris recognition [21]. In this system, images of the iris were captured and done pre-processing, and then iris feature extraction was done and compared with database images of iris. Wagh *et al.* developed a system for automatic attendance marking in 2015 based on Eigenface and PCA [22]. In this technique, the image of the whole classroom was taken from a particular angle. The captured image was enhanced by the histogram equalization technique to improve the quality of the image. The enhanced image was input for the face detection algorithm which was the Ada-Boost algorithm. Recently, Suri *et al.* developed a similar system for hospital automation using face detection and recognition algorithm [23]. The technique was useful to fill the automatic hospital admission form and sending the SMS to the concerned authority. The current study is similar to work presented by Suri *et al.* [23], but there is a difference in algorithm and implementation. Suri *et al.* [23] have used a traditional object recognition algorithm namely, Viola-Jones, which is not so accurate for tilted or turned faces [23]. Face recognition is also sensitive to lighting conditions which can damage the captured image quality. Whereas the current work proposed a facial vector-based algorithm, which eliminated the effect of lighting from the image. The proposed work also offers major advantages like, it provided manual intervention during drug delivery to the patient. Though the system is automated, it does not deliver any drug to the patient by just looking at the past medical data and act as a support system. The study [23] included only 8 dummy images (Printed newspaper cutouts) for testing and validating the system whereas, the proposed work involved roughly 200k images for testing the algorithm. Apart from this, current work has been also validated on 51 real-time images of patients. The current work captured the video of the patient and then extracted the frames from that video. This allowed us to extract as much information from the video as needed, whereas, Suri *et al.* [23] have taken static images which restrict the total available information.



**Fig. 3** Circuit diagram of the proposed system. ATmega8 was connected to 16 MHz crystal and pulled-up with  $10\text{ K}\Omega$  resistor. IC 7805 with  $100\ \mu\text{F}$  was used to generate 5 V supply. FT232 connects USB to ATmega8. Sim300 was connected for sending SMS.

### 3 Methodology

#### 3.1 Components and Equipments:

The ATmega8 IC was purchased from Mouser Electronics, India. The GSM module SIM300 was acquired from ORKA technologies, India. PIR motion sensor (Part No. 30121) was purchased from Indias Heart, India. A Logitech webcam was used. USB to serial converter FT232 (Part No. FTDIBB-F) was ordered from Robo India. 7805 voltage regulator IC was brought from Electrobot, India. Resistors, capacitors, connecting wires, jumper wires were brought locally.

#### 3.2 System design

The proposed system was designed for filling the patient admission form automatically in emergency and critical conditions. This form filled includes a detailed medical history (if available). The system comes with one more advantage of sending SMS to the concerned authority. As shown in fig. 2, image acquisition was the primary stage of the system. The USB web camera worked as an image acquisition source. The proposed system consisted of the special-case entrance (a dark passage) at the entrance of the hospital. The patients for which automated form filling was needed were carried on a stretcher through that passage. This dark passage ceiling consisted of a USB web camera with white LED lights to scan the face of the patient. The frame was extracted from captured video based on edges. The motion of the stretcher causes blurring, and hence stable frames had more edges compared to blurred frames from the video.

### 3.3 Facial vector-based algorithm

The system works with real-time video input. From this recorded video, the difference between respective frames (intra-frame difference) is computed. The maximally stable frames occur, whenever the intra-frame difference is zero. Canny edge detector was then applied to all these stable frames and the aggregated sum of all white pixels was calculated. Once the aggregated sums for all frame has been calculated, it compares with all other stable frame and the frame with the maximum aggregated sum has been selected. The frame with the maximum aggregated sum has maximum details according to edges. This frame is then selected as an input to our face detection system. As the Viola-Jones algorithm work with gray-scale images, we converted the colored image to a gray-scale image. Viola-Jones use Haar-like features to detect a face from a given image, in which all detected parts of the face are rescaled to square-shaped function.[24]. In the following step after face detection, the facial vectors are generated by tracking down relevant facial points. Face vectors are directly related to face dimensions. The face vector is nothing but values of normalized Euclidian distance between each face part. These face parts include the nose, lip, ears, forehead, eyes, etc. All face parts (landmarks) are marked using edge detection and nearest points clusters are mapped. For face vector extraction, every single pixel in the image was compared with its surrounding pixels. The change in the corresponding neighbor pixels was computed and the direction of change is noted based on the sign of the difference. The entire step is repeated for every pixel to generate a gradient map. The gradient map helped us to determine the change in the brightness and the higher the level more is the information, further, the entire image divided into 16\*16 block and find gradient block in each measured direction. Then facial landmarks are estimated so that eyes, mouth, nose and, ears, etc. are determined. These facial landmarks were resulted from the standard face and compared with the point cloud gradient that we found from our face. A point cloud is a set of data points in space. The points represent a 2D shape or object. Each point has its set of X, Y coordinates and intensity. A total of 80 specific points were mapped onto the face using the HOG pattern. HOG is defined as a histogram of gradients obtained by plotting the frequency of different gradient in an image. These points are as shown in figure 5 (d). If a particular point is present within landmark face points, we mark it by logic 1 (marked in white), else logic 0 (marked in red) gave a total 80-bit (10 bytes) face vector as shown in table 1.

### 3.4 Video frame processing for face extraction

We have applied canny edge detection and then normalized summation to extract a keyframe from the video for further steps. The image was enhanced to improve quality. After the image enhancement step, the colored image was converted into a gray-scale image. The gray-scaled image was further equalized using the histogram equalization technique. Filter was used to eliminate the noise from the image. After all pre-processing steps, face detection was carried out. Then for the face detection algorithm, all the possible junctions of edges such as nose, eyes, ears, chin, endpoints of lips, etc. were considered to form a face vector. (as shown in table 1).

Nose (px)	LY (px)	RY (px)	LE (px)	RE (px)
(400,200)	(175,125)	(175,275)	(175,30)	(175,370)
CH (px)	LP (px)	RP (px)	Face (px) height	Face (px) width
(550,200)	(475,125)	(475,275)	600	400

**Table 1** Face vector distance. The mean distances of various face features from the origin (0,0) (as shown in figure 4(k)) have been measured. The distance measured is in pixel (px) unit. Each pixel is represented by its corresponding X and Y coordinates. Here various features like eyes, lips, ears, and chin are represented by Y, P, E, and H, respectively. The positions like right left and center are represented by R, L, and C, respectively.

All the oval-shaped objects in the extracted frame were detected using Hough transform, and then the aspect ratio for each such oval-shaped objects was computed. In general, aspect ratio is defined as the ratio of the major axis to the minor axis. Oval shaped objects are the object in which both the axis (major and minor) are of different lengths. For a face (1/aspect ratio) of the rounded object lies in the range of 0.5 to 0.7.

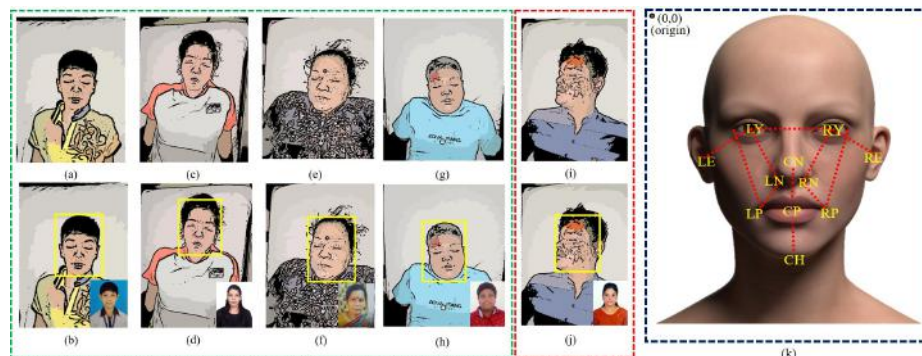
$$Aspect\ ratio = \frac{Major\ axis}{Minor\ axis} \quad (1)$$

Following the aspect ratio in the range mentioned above for all the circles in an image, the possible face was determined. The possible face was confirmed with an area-based filter. Too small circles and large circles were neglected. For our webcam resolution, the average face area was about 600 to 1200 pixels. All the face vector points are then positioned on the face, using k-means clustering and minimum distance decoding. All the missing points were generated, and for multiple faces, vector points laying in the same region (e.g., due to closed eyes) the average position of all those points was considered. Figure 4(k) shows the standard face. The standard face is a face with all the facial features marked on it along with their standard distances. We obtained a point cloud i.e. set of data points in the space representing X, Y coordinates, and corresponding intensity for the standard face whose distances measured by the pixels (px) unit. The various face features like eyes, nose, ears, lip, and chin were represented by Y, N, E, L, and H, respectively, and the positions like left, right, and center were represented by L, R, C respectively. All the face vector points, which were not fitting the cluster, were discarded. Table 3 shows all eight positions and their corresponding k-means cluster centers. As shown in table (1), the number in parentheses indicates X and Y coordinates. (400,200) px indicates the mean distance between nose and origin. Origin has been considered as (0,0) (as shown in figure 4(k)). The mean distance between origin and left eye (LY) was (175,125) px whereas it was (175, 275) px for the right eye (RY). The mean distance from the origin to left ear (LE) and right ear (RE) were (175,30) px and (175, 370) px respectively. Chin (CH) was (550, 200) px away from the origin. The mean distance from the origin to the left lip (LP) and right lip (RP) were (475, 125) and (475, 275) px. The mean height and width were 600 px and 400 px, respectively. After the face detection has been done from the image, it was cropped. This cropped image of the patient face was compared with database images. If the final face vectors of database images were found to be within the margin of 1% from the search final face vector, then 'match



found' was declared. If the confirmation is received after comparison, then the hospital admission form was filled automatically.

Automatic admission form filling involved filling general and medical history mentioned in the database into the form. The doctor can treat or medicate the patient by considering this past medical history, but during such action, manual supervision is provided on the software front end. Doctors or related medical staff can be able to change the suggested person's identity if there is a false result. This will ensure minimizing or avoiding the wrong treatment. Once the form filling has been done, SMS is sent to the nearest police station and relatives (provided their number is available in the database).



**Fig. 4** System inputs-outputs and standard face. (a,c,e,g,i) keyframes extracted from video while stretcher is passing through the dark passage. (b,d,f,h,j) Detected face region marked in the yellow rectangular box. Snippets of (b,d,f,h,j) shows matching image extracted from database. All cases, except (i) gave expected accurate results. In the (i) image, the patient face was not facing the camera, and a large wound on the face caused a mismatch with the database. Snippet (k) standard face. The face vectors were marked on the background-removed face. Here, lip (P), nose (N), eyes (Y) and ears (E) were marked using edge detection and nearest cluster analysis. The positions like right, left and center are represented by R, L, and C respectively.

To test the system, special permissions were taken from the Institute Ethics Committee (IEC) and the Medical regulatory authority of Karve hospital. All the images are attached in fig. 4 are with written authorization from concerned patients after they were recovered. The system was validated on 51 patients from Karve hospital, Thane. The system does not interrupt the usual routine of the hospital, so hospital authorities permitted us to install a special-case entrance.

### 3.5 Circuit details

The fig. 3 indicates the circuit diagram used to build the system. To implement the proposed system, we have chosen ATmega8 as the main microcontroller. For synchronization, the controller was connected to 16MHz crystal. As shown in fig. 3 pin number 1 of ATmega8 pulled up to 5 V (VCC) via 10 K $\Omega$  resistor. The power supply was driven with a lead-acid battery (6 V, 1300 mAh). The output of this battery was then fed to regulator IC 7805. The regulated 5 V output was further

Method	Accuracy (%)	Number of images	Feature dimension
Chen <i>et al.</i> [25]	92	99k	8000
Sun <i>et al.</i> [26]	92	87k	–
Huang <i>et al.</i> [27]	92	–	2302
Simoyan <i>et al.</i> [28]	93	–	256
Berg <i>et al.</i> [29]	93	20K	5000
Chen <i>et al.</i> [30]	95	99K	2000
Cao <i>et al.</i> [31]	96	99K	2000
Taigman <i>et al.</i> [32]	97	7000K	16000
Proposed method	92	200k	80

**Table 2** comparison with different face recognition and verification methods.

Database	Accuracy (%)	Total images
UCD Color Face Image[33]	96.32	299
CelebA [34]	92	200k
LFW [35]	98.23	13k
Our dataset	94.92	512
Real-time scans	94.11	51

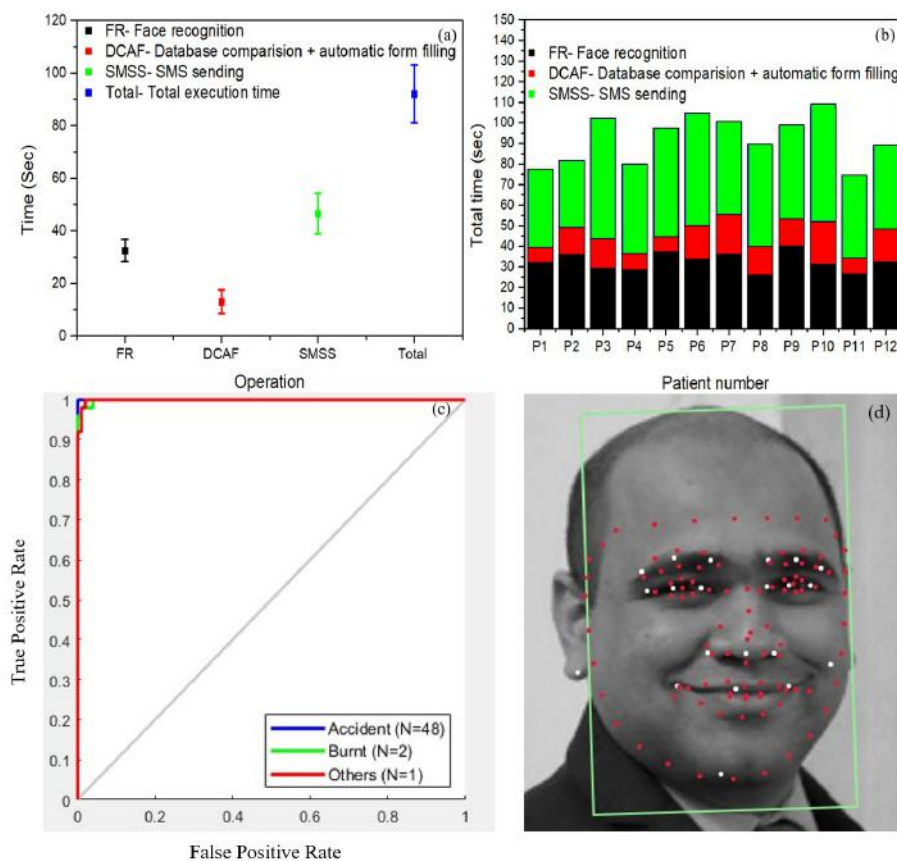
**Table 3** Accuracy of proposed method with different datasets

filtered through a 100  $\mu$ F capacitor. ATmega8 was connected to the computer using FT232. FT232 converts serial communication (i.e., RS232) signals to USB signals. The microcontroller communicates to the GSM module SIM300 through RS232 at 9600 bps. GSM module was used to send SMS to police and relatives. Since the microcontroller only had one serial communication port, the GSM module was connected through the software serial port. PIR motion sensor connected to the ATmega8 controller at pin 14.

#### 4 Results and discussions

A demonstration of the proposed system has been tested with 51 patients. The types of patients included for face verification, were varying in age, gender, and physique. To compute the results, the time delay was always considered as our output parameters since the aim of the proposed work is to reduce the delay. The total time for each operation fig. 5 (a), and total time with respect to each patient fig. 5 (b) were considered as the performance measure of the system.

The system took the overall time ( $\sim 95$  secs) (fig. 5 (a) blue) to detect and recognize the face of the patient and fill the data in the hospital form automatically. The total time taken, by the system for each patient varied (fig. 5 (b)). This variation was due to the time required to recognize the face, compare with the database, and sending the SMS. The average time taken by face detection was about  $32 \pm 4$  secs (fig. 5 (a) black). It took around  $13 \pm 4.5$  secs (fig. 5 (a) red) for comparing the image with a database and for sending the SMS it took  $46.5 \pm 7.7$  secs (fig. 5 (a) green). It is seen that the time required for sending the SMS was high and largely varying because of its dependency on network connectivity. In one case (P3) the system took almost a minute to send the SMS due to heavy rains. Due to heavy rains, there was a poor network connection to sim 300. However sending SMS is the secondary task of the system, and it does not affect the auto



**Fig. 5** Results (a) The average time taken versus each operation. Three operations that the system performs are 1) FR: patients to recognize the face (black), 2) DCAF: database comparison and automatic form filling (red) and 3) SMSS: SMS sending (green). The total time of the proposed system is the addition of these three components indicated in blue. It took around 95 secs to complete the process. (b) The bar graph of time taken by the system to recognize the face (FR), database comparison, and automatic form filling (DCAF) and SMS sending (SMSS) for the first 12 patients P1-P12. (c) Receiver Operating Characteristic Curve (ROC). It is a plot of a false-positive rate versus a true positive rate. We found that it is very close to the ideal ROC curve in three different kinds of cases (i.e. Accident, burn, and others). (d) Sample 80 point face vector.

form filling and hospitalization. In general, another issue, along with the network problem that we faced during the actual testing, was people with facial hairs. Facial hair causes extra feature vectors which makes the system slower in identifying the person. Fig. 5 (c) shows a receiver operating characteristic curve or a ROC curve. The ROC curve has been created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. In this study, we have prepared the ROC curve for both training and testing datasets. For plotting ROC for training, we had used the celebA, LFW, and UCFI database. The database included 200k celebrity images. For testing purposes, we included 51 real-time

images of patients (48 were road accident cases, 2 were burnt cases and 1 was of face stab.)

Fig. 4 shows practical inputs received (4 a,c,e,g,i) , and corresponding outputs (4 b,d,f,h,j) of the proposed system. The system successfully interpreted the result of 48 patients out of 51 (94.11% accuracy). In fig. 4 only 5 cases out of 51 are discussed. In the first four (4 a,c,e,g) cases, the system interpreted results accurately whereas in the fifth case(4 i), the system was unable to detect the patient as the feature vector extraction algorithm failed due to wounds and non-fronto-parallel image. In the first case (4 a), a 12 years boy met an accident and brought to the hospital when he was unconscious. He was passed through the system. The keyframe image (4 a) was extracted from the real-time video captured. This keyframe is used to detect a face, as shown in fig. 4 b. The snippet indicates the matched database image. The system accurately detected the identity of the patient and also fill the form. The system also sent the message to the parents of the patients, who could be unaware of the incident. The second case (fig. 4 c) was a 25-year old female with a cut in the vein, was also detected correctly even with closed eyes. The third case (fig. 4 e), 58 years old lady got admitted, the system took approximately 0.4 seconds extra to recognize the face because of the head bindi. The 15 years boy (fig. 4 g) with head injury also gave a proper match. All the extra feature vectors appeared due to blood and wound were discarded during the k-means stage. We have also reported one failed case here for future researchers. In this case(fig. 4 i), 30 years man was brought to the hospital who had met with the road accident. The face was wounded, and the angle at which the camera captures was not fronto-parallel hence matching algorithm failed. We also included 2 patients from burnt cases, where the face was 20 % to 35 % damaged. The system was able to correctly recognizes these patients also. The average time taken by the system to recognize the person and fill the form was about 2.5 and 1.8 minutes respectively. Among all the admitted cases three cases were failed to identify because of extensive blood on the face. At the same time, we found that the occlusions like bindi and facial hair increased the processing time slightly (less than 0.5 sec), as it increased the number of points to be processed.

The system is meant only to fill the form during the hospital admission in emergency cases, so as it will reduce the further delay for treatment. Although the procedure to fill the form is completely automatic, we have made the provision for manual supervision in the software. There are chances when false matches may occur, in such cases, if drug delivery has been suggested for the patient based on past history then it could be disastrous. Manual supervision provided to hospital staff will help to overcome this issue. By manual supervision, we can avoid any mishaps during treatment and also can avoid false drug entry in the patient's body.

There could be cases of incorrectly identifying someone when using facial recognition and can cause a catastrophe. In order to avoid it fully or to minimize such situations, we have provided a manual supervision facility at least before starting drugs based on past medical history. Sending wrong SMS to the wrong family is still a downside of our system.

#### 4.1 Comparison with other methods and datasets

To compare the results of the proposed method with different datasets, we used the celebA [34], LFW [35], and UCFI [33] datasets. We also created our own dataset with 512 images. As shown in table 3, the celebA [34] dataset consisted of 200k images whereas LFW [35] and UCFI [33] consisted of 13k and 299 images respectively. We found that the highest accuracy of 98.23 % was achieved from the LFW [35] dataset. The lowest accuracy of 92 % was achieved with the CelebA [34] dataset. The dataset created by us with 512 images, gave an accuracy of 94.9 %. Also, a real-time scan for 51 patients gave us an accuracy of 94.11%. As shown in table 2, the performance of the system was compared with various existing methods. The accuracy of our proposed method was comparable to the other different methods reported in the literature. Taigman *et al.* [32] achieved maximum accuracy of 97 % also their number of images was the highest. But on the other hand, their method required a 16k long feature vector compared to the proposed method of the 80-bit vector. Our feature vector is the smallest still the accuracy is within the range of one reported in the literature. The small size feature vector and simplicity make the proposed algorithm more suitable to implement on the portable systems.

### 5 Conclusions

Avoiding delay in starting treatment could save a lot of lives. We proposed a system that provided the solution for automatic form filling and sending the SMS to police and relative automatically and thus avoiding delays. The automated form filling was achieved by face detection. The face detection was based on a novel method of calculating the face vectors. The system was tested with 51 patients in emergency conditions, and we found that the system worked with an accuracy of almost 94.11%. The system takes a relatively large time to calculate the face vectors in special cases like bindi on the forehead and facial hair. We also recommend using the proposed system with all the time-critical processes, such as interconnecting flights at international airports. Another possible application of this system is for people with disabilities, such as people who cannot speak or deaf.

### 6 Acknowledgements

First of all, We would like to express my sincere gratitude to the management, doctors, and supporting staff of Karve Hospital, without their support, we could not have verified the results in the field. Secondly, We would like to thanks all the 51 anonymous patients of Karve Hospital for allowing us to used their photos in the manuscript for a better understanding of our work. We are grateful to Mr. Ghone (stretcher boy) who helped us every time standing at the entrance of the special- case entry to pass the patient through the special case entrance. Finally, we would like to thank all the colleagues of Ninad's research lab and K. J. Somaiya college of engineering who made this work possible.

## Compliance with Ethical Standards

### Conflicts of interest

Authors M. Parab and N. Mehendale declare that they have no conflict of interest.

### Involvement of human participant and animals

This article does not contain any studies with animals or Humans performed by any of the authors. All the necessary permissions were obtained from the Institute Ethical Committee and concerned authorities to use video captures of patients.

### Information about informed consent

Informed consent was acquired from all human participants whose videos were used for this Novel work.

### Funding information

No funding was involved in the present work.

## References

1. P. Souminen, A. Kivioja, J. Öhman, R. Korpela, R. Rintala, K.T. Olkkola, Severe and fatal childhood trauma, *Injury* **29**(6), 425 (1998)
2. W.H. Organization, et al., Global status report on road safety 2018 (2018)
3. NDTV, et al. Road accident statistics in india (2016)
4. L. Tran, X. Yin, X. Liu, in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2017), pp. 1415–1424
5. V. Singh, L. Malkunje, S. Mohammad, N. Singh, S. Dhasmana, S.K. Das, The maxillofacial injuries: A study, *National journal of maxillofacial surgery* **3**(2), 166 (2012)
6. A.M. Black, D.A. Patton, P.H. Eliason, C.A. Emery, Prevention of sport-related facial injuries, *Clinics in sports medicine* **36**(2), 257 (2017)
7. G. Canzi, E. De Ponti, G. Novelli, F. Mazzoleni, O. Chiara, A. Bozzetti, D. Sozzi, The cfi score: Validation of a new comprehensive severity scoring system for facial injuries, *Journal of Cranio-Maxillofacial Surgery* **47**(3), 377 (2019)
8. M. Trokielewicz, A. Czajka, P. Maciejewicz, in *2016 International Conference on Biometrics (ICB)* (IEEE, 2016), pp. 1–6
9. T.A. Coulibaly, R. Béogo, I. Traoré, H.M. Kohoun, B.V. Ili, Inter personal violence-related facial injuries: a 10-year survey, *Journal of Oral Medicine and Oral Surgery* **24**(1), 2 (2018)
10. P. Majumdar, S. Chhabra, R. Singh, M. Vatsa, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (2018), pp. 2173–2179
11. P. Majumdar, S. Chhabra, R. Singh, M. Vatsa, Subclass contrastive loss for injured face recognition, arXiv preprint arXiv:2008.01993 (2020)
12. J. McDonagh, G. Tzimiropoulos, in *European Conference on Computer Vision* (Springer, 2016), pp. 569–580
13. K.C. Yow, R. Cipolla, Feature-based human face detection, *Image and vision computing* **15**(9), 713 (1997)
14. B. Jun, D. Kim, Robust face detection using local gradient patterns and evidence accumulation, *Pattern Recognition* **45**(9), 3304 (2012)

15. J. Mahajan, A. Paithane, in *2017 International Conference on Inventive Communication and Computational Technologies (ICICCT)* (IEEE, 2017), pp. 439–444
16. N. Mehendale, Facial emotion recognition using convolutional neural networks (ferc), *SN Applied Sciences* **2**(3), 1 (2020)
17. H.A. Rowley, S. Baluja, T. Kanade, in *Advances in Neural Information Processing Systems* (1996), pp. 875–881
18. T.K. Leung, M.C. Burl, P. Perona, Finding faces in cluttered scenes using random labeled graph matching (1995)
19. G. Yang, T.S. Huang, Human face detection in a complex background, *Pattern recognition* **27**(1), 53 (1994)
20. N.T. Deshpande, S. Ravishankar, Face detection and recognition using viola-jones algorithm and fusion of pca and ann, *Advances in Computational Sciences and Technology* **10**(5), 1173 (2017)
21. S. Kadry, M. Smaili, Wireless attendance management system based on iris recognition, *Scientific Research and essays* **5**(12), 1428 (2013)
22. P. Wagh, R. Thakare, J. Chaudhari, S. Patil, in *2015 International Conference on Green Computing and Internet of Things (ICGCIoT)* (IEEE, 2015), pp. 303–308
23. N. Suri, M. Marne, M. Ghotekar, U. Pacharaney, in *2016 International Conference on Inventive Computation Technologies (ICICT)*, vol. 1 (IEEE, 2016), vol. 1, pp. 1–6
24. R. Lienhart, J. Maydt, in *Proceedings. international conference on image processing*, vol. 1 (IEEE, 2002), vol. 1, pp. I–I
25. D. Chen, X. Cao, L. Wang, F. Wen, J. Sun, in *European conference on computer vision* (Springer, 2012), pp. 566–579
26. Y. Sun, X. Wang, X. Tang, in *Proceedings of the IEEE international conference on computer vision* (2013), pp. 1489–1496
27. C. Huang, S. Zhu, K. Yu, Large scale strongly supervised ensemble metric learning, with applications to face verification and retrieval, arXiv preprint arXiv:1212.6094 (2012)
28. K. Simonyan, O.M. Parkhi, A. Vedaldi, A. Zisserman, in *BMVC*, vol. 2 (2013), vol. 2, p. 4
29. T. Berg, P.N. Belhumeur, in *Bmvc*, vol. 2 (Citeseer, 2012), vol. 2, p. 7
30. D. Chen, X. Cao, F. Wen, J. Sun, in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2013), pp. 3025–3032
31. X. Cao, D. Wipf, F. Wen, G. Duan, J. Sun, in *Proceedings of the IEEE international conference on computer vision* (2013), pp. 3208–3215
32. Y. Taigman, M. Yang, M. Ranzato, L. Wolf, in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2014), pp. 1701–1708
33. P. Sharma, R.B. Reilly, in *Proceedings EC-VIP-MC 2003. 4th EURASIP Conference focused on Video/Image Processing and Multimedia Communications (IEEE Cat. No. 03EX667)*, vol. 1 (IEEE, 2003), vol. 1, pp. 423–428
34. Z. Liu, P. Luo, X. Wang, X. Tang, Large-scale celebfaces attributes (celeba) dataset, Retrieved August **15**, 2018 (2018)
35. A. Jalal, U. Tariq, in *Asian Conference on Computer Vision* (Springer, 2016), pp. 531–540