National University of Computer & Emerging Sciences (FAST-NU)

EyeHope

HELPING BLIND PEOPLE GET A LITTLE MORE INDEPENDENT

Project Supervisor: Dr. Zulfiqar Memon

Project Team

Hammad Mubarak	**k132113**
Aamir Amin	**k132072**
Mahzain Malik	**k132004**
Saman Karim	**k132071**

Submission Date
28 April, 2017

Submitted in partial fulfillment of the requirements for the degree of Bachelor of Science

The Department of Computer Science
National University of Computer & Emerging Sciences (FAST-NU)
Main Campus, Karachi
May 2017

National University of Computer & Emerging Sciences (FAST-NU)

EyeHope

HELPING BLIND PEOPLE GET A LITTLE MORE INDEPENDENT

Project Supervisor	**Dr. Zulfiqar Memon**
Project Manager	
Project Team	**Hammad Mubarak k132113** **Aamir Amin k132072** **Mahzain Malik k132004** **Saman Karim k132071**
Submission Date	**28/April/2017**

Supervisor: Dr. Zulfiqar Memon _____

Head of Department: Dr. Jawwad A. Shamsi _____

The Department of Computer Science
National University of Computer & Emerging Sciences (FAST-NU)
Main Campus, Karachi

Acknowledgement

We are highly grateful to Dr. Zulfiqar Memom (Project Supervisor), Mr. Sayed Yousuf (Project Co-Supervisor), Mr. Waqar Chaudry (Director, DEO consultancy, London), Mr. Waqar Puri (Projects coordinator and administrator at Special Talent Exchange Program "STEP") and Mr. Nauman Atique for their precious guidance and constant supervision as well as for providing us necessary information regarding the project and also for their support in completing the project. We would like to express our gratitude towards our parents and member's of NUCES-FAST University for their kind co-operation and encouragement which helped us in the completion of this project. We would like to express our special gratitude and thanks to STEP Organization for giving us such attention and time. Our thanks and appreciation also goes to our entire colleague's in developing the project and to all those people who willingly helped us for the successful completion of this project.

Document Information

Category	Information
Customer	NUCES-FAST
Project Title	EyeHope
Document	FYP I Final Report
Document Version	1.0
Identifier	ZM014 Final Report
Status	Final
Author(s)	Hammad Mubarak, Aamir Amin, Mahzain Malik, Saman Karim
Approver(s)	Dr. Zulfiqar Memon
Issue Date	

Definition of Terms, Acronyms and Abbreviations

Term	Description
DFD	Data Flow Diagram
STEP	Special Talent Exchange Program

Table of Content

1. CONTEXT AND PRELIMINARY INVESTIGATION

1.1 Project Selection

Advanced technology is an integral part of the twenty first century. The fascinating ideas of upcoming scientists are reviving the lifestyle of those who are physically well and want to make their life even more convenient. But what about those who are already bereaved their physical resources due to any reasons? Specifically for those who are blind or visually impaired, this application will make them visualize some of the components of a specific environment/area in an efficient manner, based on the real-time computer vision through OpenCV on android platform. This technological initiative will pull them out from the sphere of white sticks/canes and will guide them towards a self-reliant and independent world. The motive behind selecting this kind of project is to use the current advanced technological resources to help someone live their life conveniently and independently. We thought that we can utilize our abilities and technical skills to enhance the lifestyle of blind and visually impaired people. As stated above, our first and the only priority were to do something that could be supportive and helpful for the human kind in future.

Another reason for selecting this project was that, there are many organizations that encourage and support these kinds of projects which are specifically made for disabled people. Our team is having the support of Special Talent Exchange Program (STEP) for this project. They will not only mentor us, but provide us a platform to promote and market it among the disabled community. This project will somehow act as a sight sense for blind and visually impaired users and will facilitate them to communicate conveniently with others.

1.2 Project Background

Human senses are an essential part of human body through which a person can manage to survive in the current society. But in the case when a person is blind or visually impaired, the inability to identify the presence of people normally at home or during meetings becomes inconvenient for them [1]. Real-time face recognition [4], text recognition [2] and object detection [5] are some of the dominant developed applications. Machine learning employs algorithms that can find patterns from exemplars and make data driven predictions. Thus, computers have the ability to learn and act without being given explicit directions by mimicking the human cognitive framework of collecting and applying knowledge to make decisions [7]. Furthermore, algorithms used for face detection are proposed by Voila and Jones [4] and Eigenfaces (Principle Component Analysis) are used in order to identify gestures, facial expressions and emotions [3].

The present work is focused on developing face detection and face recognition algorithms to be used by visually impaired people [1]. The previous research also says that body movement can also become a part for detecting emotions of the humans. But the emotion communication through bodily expressions has been a neglected area for much of the emotion research history [6]. Over the last two decades, researchers have significantly advanced human facial emotion recognition with computer vision techniques. Historically, there have been many approaches to this problem, including using pyramid histograms of gradients (PHOG) [8], AU aware facial features [9], boosted LBP descriptors [10], and RNNs [11]. However, recent top submissions [12], [13] to the 2015 Emotions in the Wild (EmotiW 2015) contest for static images all used deep convolutional neural networks (CNNs). The classification of emotions is also a huge milestone to cover. As, In classification problems, good accuracy in classification is the primary concern; however, the identification of the attributes (or features) having the largest separation power is also of interest [14].

To be specific, computers are pushing the limits to reach what seem to be singularly human capabilities, image and emotion detection are at the forefront of scientific efforts [7].

1.3 Literature Review

Understanding emotional facial expressions accurately is one of the determinants in the quality of interpersonal relationships. The more one reads another's emotions correctly, the more one is included to such interactions. The problems in social interactions are shown in some psychopathological disorders may be partly related to difficulties in the recognition of facial expressions. Such deficits have been demonstrated in various clinical populations. Nonetheless, with respect to facial expressions, there have been discrepant findings of the studies so far [15]. The process of emotion recognition involves the processing images and detecting the face then extracting the facial feature. Facial Expression Recognition consists of three main steps. In first step face image is acquired and detect the face region from the images and pre-processes the input image to obtain image that have a normalized size or intensity. Next is expression features are extracted from the observed facial image or image sequence. Then extracted features are given to the classifier and classifier provides the recognized expression as output [16]. In the past researches, it is also stated that Principal Component Analysis [17], Local Binary Pattern (LBP)[18],Fisher's Linear Discriminator[19] based approaches are the main categories of the approaches available for feature extraction and emotion recognition [16]. There are various descriptors and techniques used in facial expression recognition like the Gradient faces, local features, local binary pattern (LBP), local ternary pattern(LTP), local directional pattern (LDiP)and Local derivative pattern (LDeP) [20].

Many studies on facial expression recognition and analysis have been carried out for a long time because facial expressions play an important role in natural human-computer interaction as one of many different types of nonverbal communication cue [21]. Paul Ekman et al. postulated six universal emotions (anger, disgust, fear, happiness, sadness, and surprise), and developed Facial Action Coding System (FACS) for taxonomy of facial expressions [22].

1.4 Economic Feasibility

Generally, the advance technological project requires a large amount of budget to make its roots more strong and provide desired output. But some are those which require fewer budgets and provide more reliable output/services. Economic feasibility of a project consists of many aspects such as target customers, product specifications, assets, manpower and etc. EyeHope is perfectly suitable in all these aspects.

As, target customers of EyeHope are focused and determined. These are none other than the blind and visually impaired people who curiously requires such kind of application that can assist them to understand the expressions and emotions of society which exist near to them. The product specifications of EyeHope application are that it mainly focuses on the count of number of people in front of the user and their emotions. Usage of technological tools and software's is also one of the important aspects of economic feasibility. The EyeHope is an android application which requires OpenCV for its large number of functionalities. For executing and using this application, a user will require an android smart phone with the feature of camera and accessibility mode. The application does not require any network or internet assistance.

This application is developed by some developers under the supervision of some professionals due to which it does not require a large number of manpower for its construction and maintenance.

1.5 Project Scope

Human and computer interaction is advancing step by step. More up to date advances are created oftentimes and more current methods for conveying and interfacing with a PC are presented. From keypad contributions to touch contributions to hand signals, innovation has upset the way a client interfaces with a gadget. To make it even a stride advance, designers are taking a shot at innovations which make it conceivable to connect with gadgets utilizing simply outward appearances, whatever the utilization. Visually challenged people face difficulty while interacting with others. This android based application will overcome the difficulty faced during interaction and will help them to observe one's emotions with some other important features. The scope of this project is to provide something worthy that can help the targeted audience to increase the communication level with other people and to make them independent from any kind of support. The scope of this project also cares about cost efficiency and affordances, as the application will be developed under a specified budget which will be affordable for the user. The application will detect the faces which are in front of camera, communicate the number of people in front of camera to the user and finally will detect and communicate the emotions, of the detected faces, to the user.

1.6 Project Objectives

The way of communication with others is a significant part of one's personality. Visually challenged people face difficulty while interacting with others. The basic objective behind this application is to overcome the difficulty faced during interaction and to decrease the communication gap between blind/visually impaired people and others. The technique behind the application will assist blind/visually impaired users to perceive and visualize the number of people in front of them and expressions and emotions of detected people in a specific environment. Specifically, the EyeHope application will be useful for all those visually challenged people who face number of fundamental communicational challenges in current society.

1.7 Deliverables

The application is previously divided into certain components/modules. After merging these components, the final deliverable will be the full flash android application on Google Play Store, named "EyeHope (Helping blind people get a little more independent)". The deliverable will also include a proper user guide with instructions properly specified. It will also include designs and description of application in a proper documented form for technical and other purposes.

2. RESEARCH

2.1 Primary Research

There is always a lot of research, human and technological efforts behind every successful project. Doing research for a particular project makes a project stronger on the basis of its design and implementation. The EyeHope application also contains a long research history, as the main components of its primary research are its users, environment in which it will be used and etc. It was observed that the fascinating ideas of upcoming scientists are reviving the lifestyle of those who are physically well and want to make their life even more convenient. But what about those who have already bereaved their physical senses due to any reasons? Among these physically challenged people, the main focus was on blind/visually impaired people as they face more difficult challenges than any other. It was observed that this group of people face difficulty while communication with others. Many of blind/visually impaired users were interviewed by EyeHope application team. According to the answers given by these users, the points which were catchy and to be focused more were that when the people in front of these blind/visually impaired users are talking, senses of these users are attentive and they can easily understand the gestures of the person in front of them through the tone of their voice and the way of talking. The moment the person in front of user becomes quite, sensors of users stop working (in terms of understanding what is the person in front of the user is doing. what are his/her actions? What gestures/emotions/expressions the person has currently on their face? Are they still present in front of the user or not?). This piece of information was kept in the mind of application's planning team and convinced them to build something innovative that can help these users to communicate with the people in front of them conveniently.

2.2 Academic Research

2.2.1 Development Tools

2.2.1.1 Comparison of Different Tools

EyeHope application is a type of computer vision based application which detects the number of people in front of the application device and also detects the emotions of already detected people in real time environment. These milestones can be achieved through different developmental tools such as android studio (with OpenCV), Matlab software's and etc. If we compare these developmental tools, on the basis of speed, Matlab is built on Java, and Java is built upon C. So when you run a Matlab program, your computer is busy trying to interpret all that Matlab code. Then it turns it into Java, and then finally executes the code. OpenCV, on the other hand, is basically a library of functions written in C/C++. You are closer to directly provide machine language code to the computer to get executed. So ultimately you get more image processing done for your computers processing cycles, and not more interpreting. If these tools are compared on the basis of the resources it needs than, Due to the high level nature of Matlab, it uses a lot of your systems resources. Matlab code requires over a gig of RAM to run. In comparison, typical OpenCV programs only require ~70mb of RAM to run in real-time.

2.2.1.2 Selection of Appropriate Tools

As it has been discussed above, the development environment which best suits the requirements of this application is android studio with OpenCV. As it is a smart phone application, the android was specifically preferred because majority of the users contains android operating system based smart phones.

2.2.1.3 Development Tool Reasoning

Android (with OpenCV) is specifically use because programs written in OpenCV run much faster than similar programs written on other platforms. OpenCV is must faster when it comes to speed of execution. In terms of memory or storage; typical OpenCV programs only require ~70mb of RAM to run in real-time [23]. In terms of cost, the OpenCV (BSD license) is free. Finally, in terms of portability, any device can run C. Android (with OpenCV) is a memory optimized solution and runs more efficiently than other platforms [24].

2.2.2 Secondary Research

For EyeHope application, Web url's and research papers related to computer vision has played a important role in understanding and selection of most suitable development tools and software for the construction of application. Different links and repositories are used for the collection of training and testing dataset consists of images of different emotions. Android and OpenCV documentations were also used to develop the application in an efficient and professional way.

3. REQUIREMENT ANALYSIS

3.1 User Requirements/Use Cases

There are several requirements which are required by the blind and visually impaired user's to use the EyeHope application in a convenient way. The final designed use cases are as follows,

1) Tap to start detection

2) Tap to stop detection

3) Navigate through application

4) Use Time interval option

5) Use on request option

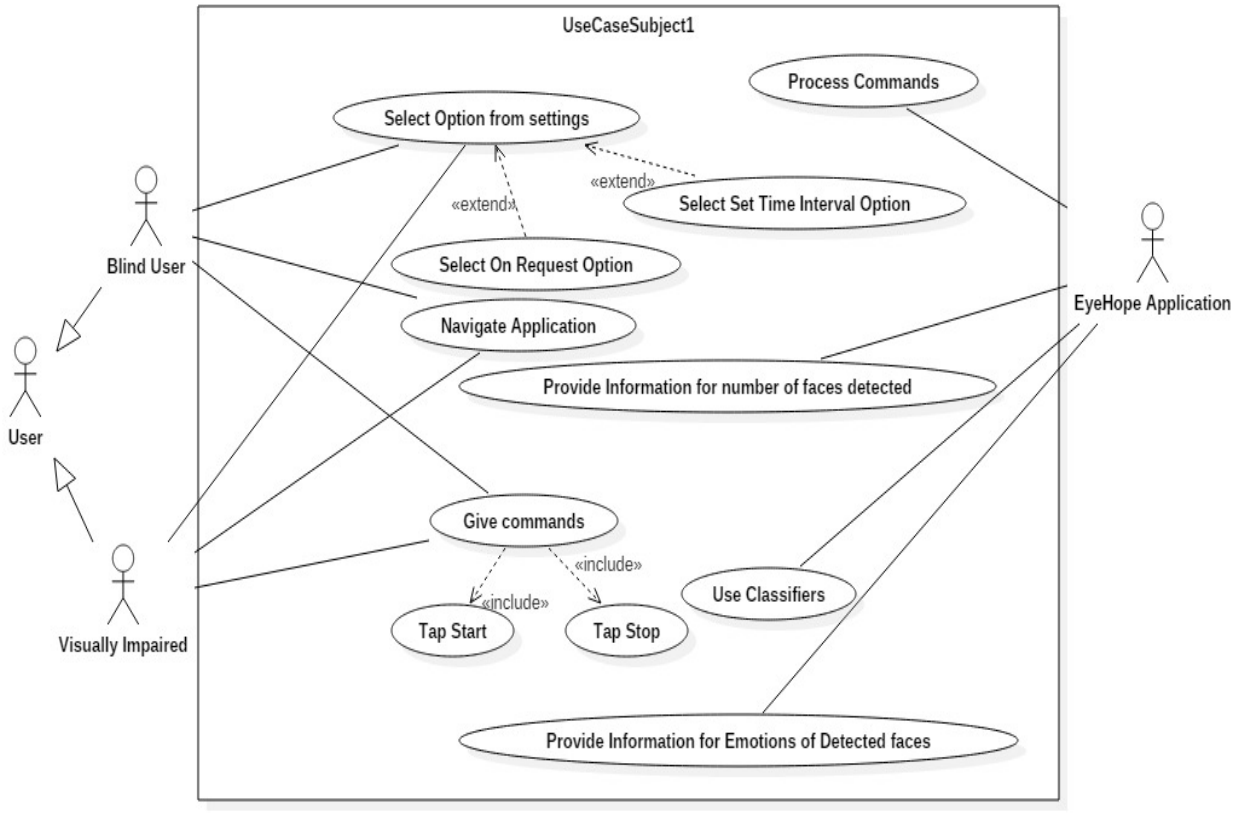
6) Use filter emotions option

7) Use Help option

8) Use change face size option

3.2 Use-Case Diagram

In the use case diagram of EyeHope, there are three types of actors which are directly related to EyeHope application. The actors from user's side are blind user and visually impaired user. From application side, EyeHope application is the third actor. The blind user's are related to several use cases such as select options from settings window, navigate towards application and give commands by tapping on desired options. The visually impaired user has same use cases as of blind users. The blind and visually impaired actors can also use change face size option to change the frame size which shows the detected faces. The EyeHope application also allows user to use the filter emotions option from which the user will get updates for only desired/selected emotions. A user can also use help option to get more information about the application. The user can also use on request option, from which user can get output only when he/she wants. The EyeHope application is also related to several use cases which are, processing the commands given by user provide information for number of faces detected, using classifiers and provide information for emotions of detected faces. The use case diagram for EyeHope application is shown below.



3.3 System Specification

3.3.1 Functional Requirements

There are number of functional requirements for EyeHope application which are as follows,

- The mobile application should have the option to start/stop the process of analyzing number of people and their emotions.
- The mobile application should also have option to set intervals of capturing images and also have the option for user to request a capture at the current time.
- The whole mobile application should be designed in way that it can be easily accessed and used through speech by the blind.
- The mobile application should have the HELP section.
- The HELP section should also be operated through voice and should give output in speech.
- The application should confirm the selected options from the user through voice question answering.
- On double tap gesture, the application should go on listening mode means to take input from the user in voice.
- The application should tell number of people in the image.
- The application should tell the emotions of the people in the frame.

3.3.2 Non-Functional Requirements

The non-functional requirements for EyeHope application are as follows,

3.3.2.1 Nature of Users

The EyeHope mobile application is specially designed for the kind of users which are bereaved from the sight senses (specifically blind and visually impaired people). So, blind and visually impaired users are specifically targeted and are required for this application.

3.3.2.2 Error Handling

If the processing in application reaches a threshold level or mobile phone suddenly stops giving responses, the application automatically shutdown's all its backend and frontend processes and then can again be restarted.

3.3.2.3 Performance Constraints

The machine trained model will be in the mobile application so all the processing will be done in the mobile, there is no need for any external servers which takes time so the user will get the result quicker every time. It will give the result within 5 seconds of the image captured.

3.3.3 Quality Requirements

A systems quality is the one, which takes it to the satisfaction of the targeted audience. The quality of a system make sure that whether different components of the system are convenient for the user to use and are they efficient enough to use? There are certain quality requirements which are as follows,

3.3.3.1 Maintainability

Maintainability involves a system of continues improvement, which means learning from the past in order to improve the ability to maintain the system. The application will be continuously observed by the developing team. It will be taking cared about all the unexpected breakdowns of the application, all the observed defects will be corrected as soon as possible, the features of the application will be modified or new features will be added according to the changing requirements of the targeted users.

3.3.3.2 Simplicity

The factor of simplicity is important for a reliable, error free and a successful system. Making system complex can make the users irritating and suffer badly. More complex system has less life. The relevant, required and simple system is the key to reliability and success. In EyeHope application, the interface is kept simple and user friendly for both kind of users (blind and visually impaired). The blind can clearly listen the simple Talk back output sound and the visually impaired can easily see the text and colors on the screen as the text is of large and simple font as well as the colors are simple and convenient for the users to figure it out.

3.3.3.3 Extensibility

The EyeHope application definitely contains the extensibility factor in it. As many of its modules can be modified to a certain extent for example the frontal face detection feature can be extended by detecting the face even if camera is facing only left side of the face or only right side of the face. Many new features can also be a part of this application such as text detection by processing the frames having any kind of text in it, Face Recognition, Object recognition and much more.

3.3.3.4 System Security

The system does not contain any kind of personal or confidential information of users due to which the system security factor is not much required.

3.3.4 Interface Requirements

3.3.4.1 User Interface

A Talk back accessibility mode will be required to guide the blind or visually impaired user about the navigation or path from which the user will be updated that at which screen the user currently is and on which option the finger of user is on. On a single tab, the accessibility mode will tell about the option which is currently taped by user and if the user wants to go with the same option, then he/she will double tap on the same area of the screen which will activate and will open the new window related to that selected option.

3.3.4.2 Hardware Interface

A camera that could capture the images and a pair of ear phones to facilitate the blind with the output in speech.

3.3.4.3 Software Interface

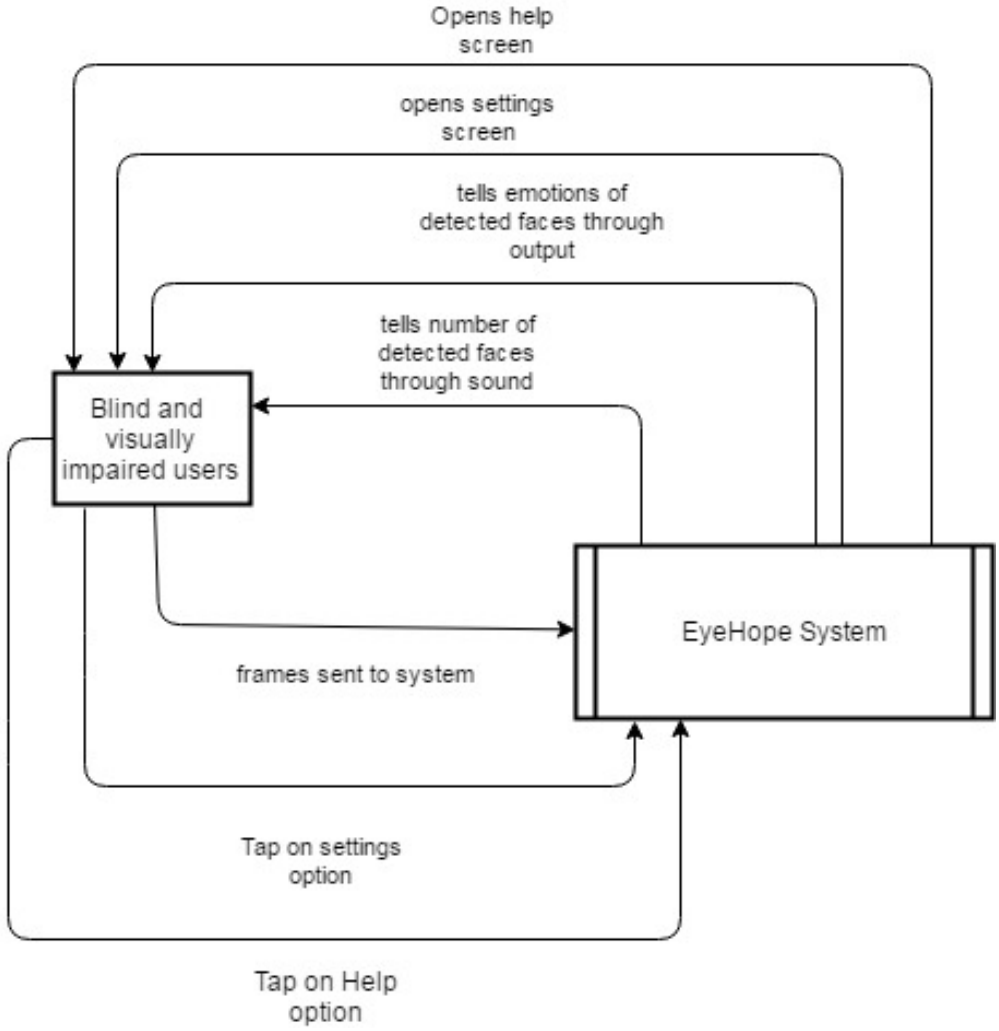
- They application will be developed on android studio 2.1 and will be supportable by the android versions 4.1 (Jelly Bean) and above.
- Open CV library is used in android studio for image processing.
- Support Vector Machine (SVM) is used for machine learning and training of data set.
- Haar Cascades are used for face detection.

4. DESIGN

4.1 Deliverables of Process Modeling

4.1.1 Context Diagram

The figure below shows the context level diagram of EyeHope application. There is a strong mode of communication between user and system. When the user double taps on camera option on the screen, the camera becomes activated and starts capturing the frames. The frames are forwarded to the backend of the system. The system provides an audible output as the number of faces detected from the frame. These frames are then also used to detect the emotions for each detected faces. A user can double tap on settings option to access and change its configurations. A user can also double tap on help option to access user manual details.

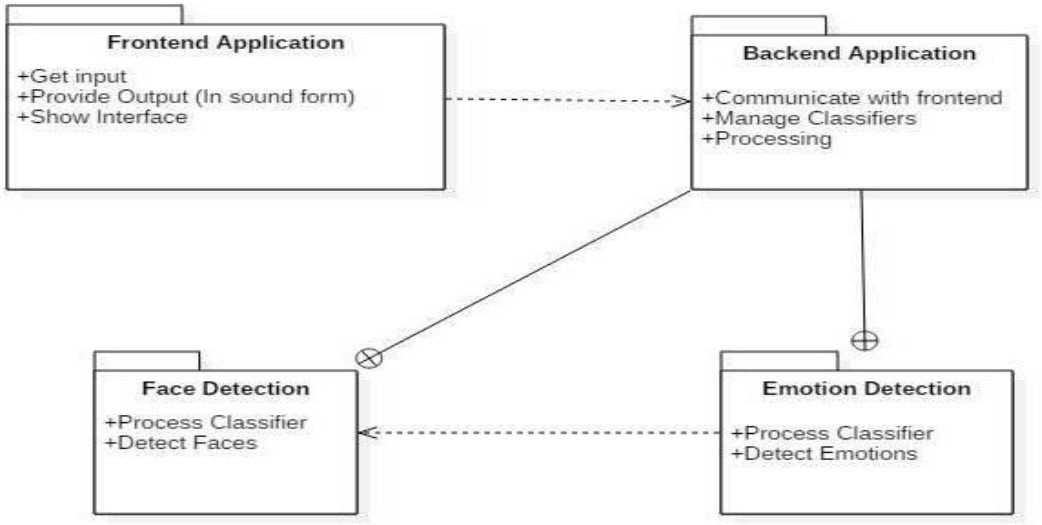


4.1.2 System Architecture

A System Architecture or System design is the theoretical as well as a diagrammatic model that characterizes the structure, conduct, and more perspectives of a framework. A design portrayal is a formal depiction and representation of a framework, composed in a way that supports thinking about the structures and practices of the framework. The EyeHope Application is a vast and dynamically behavioral android application which intelligently detects the faces which are captured by the camera and emotions are precisely detected with the help of classifiers. The application consists of different number of components/modules which contains its own functionalities and responsibilities and is also interdependent on other modules/components. The system architecture of EyeHope application is briefly and clearly described through certain aspects such as Package diagram, deployment diagram, component diagram and subcomponent diagrams. The architecture described on the next pages will explain the decomposition of the system at top level, in a way that will provide a strong foundation for detailed design work. The first part of system architecture, which will show the system divided into several packages, is shown below,

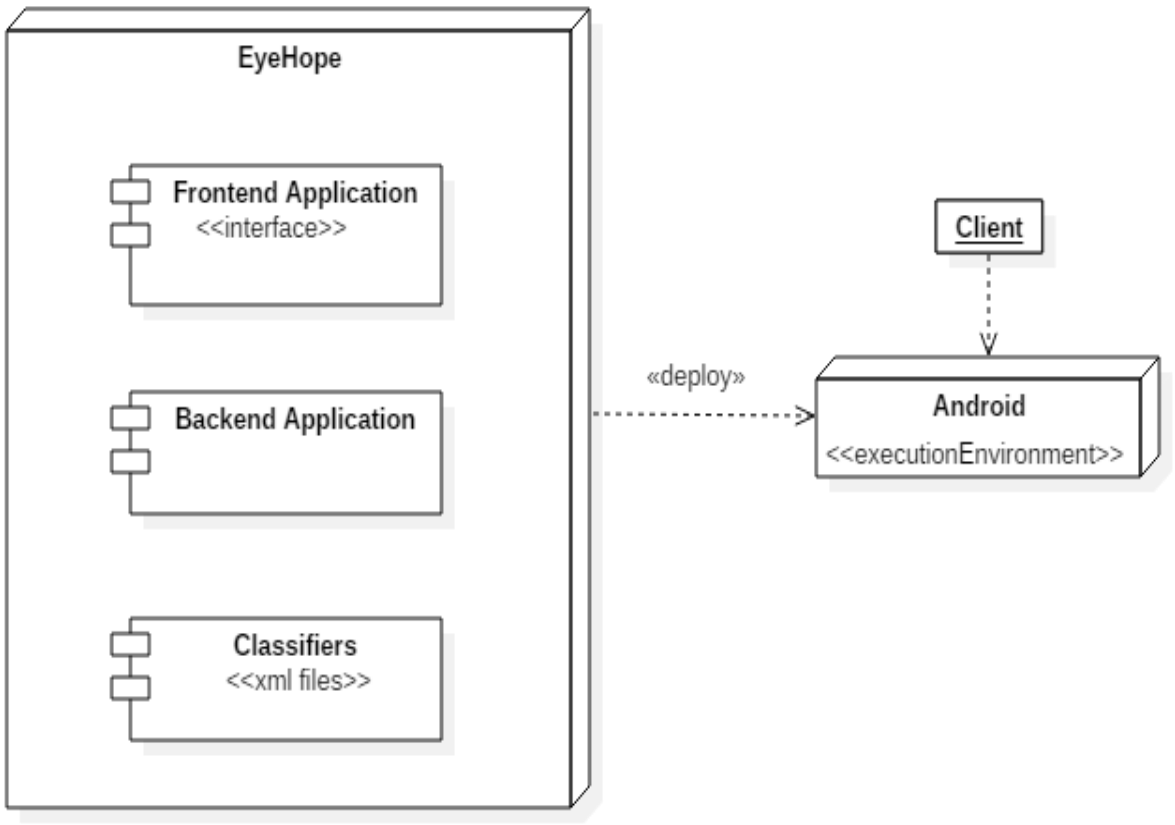
4.1.2.1 Package Diagram

A package diagram in the Unified Modeling Language delineates the conditions between the packages that make up a model. A package import is a relationship between a bringing in namespace and a package, demonstrating that the bringing in namespace includes the names of the individuals from the package to its own namespace. As mentioned above, EyeHope is a diversified application which helps more than one type of user's to live their life in a convenient way. As it is an immense application, it contains different packages which are shown in the package diagram below. One of the packages shown in below diagram is named "frontend Application" package. The frontend package contains the functionalities and responsibilities of the interface side such as "Get input" from user in any form such as through voice or by tapping on screen options. The User "Get output" from applications interfaces such as through sound. It is the responsibility of frontend application to provide a user friendly and convenient interface to its users. The functionality and responsibility of frontend application package is to accurately communicate with frontend interface. As well as it is responsible to manage and process the classifiers provided to the application. The face detection and emotion detection packages have responsibilities to access and process their respective classifiers and access backend application to detect faces and their emotions accurately.



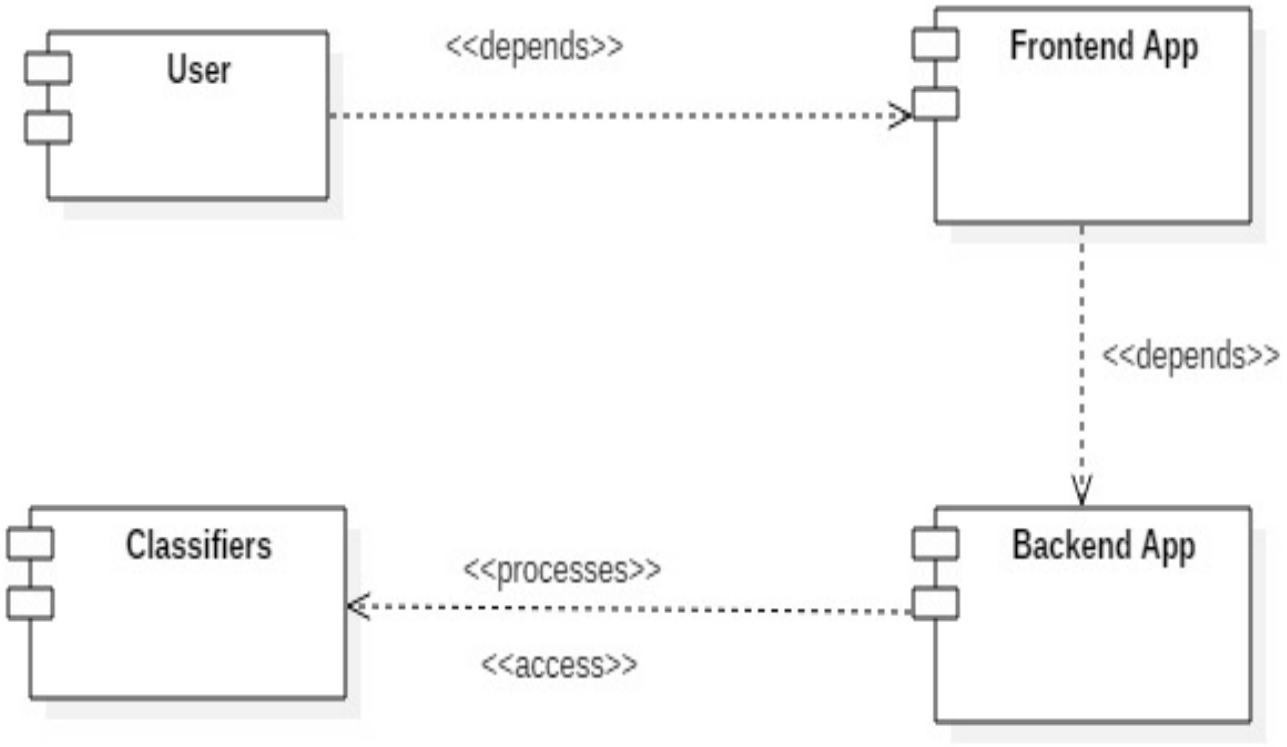
4.1.2.2 Deployment Diagram

Deployment diagram is a structure diagram which demonstrates design of the framework as sending (dispersion) of programming ancient rarities to arrangement targets. Artifacts speak to solid components in the physical world that are the result of an improvement procedure. The file shaped box in the diagram below is the package which includes multiple nodes (components).The EyeHope package contains three components which are Frontend application which is basically interface of the application. Another component is Backend application component which contains all the hidden functionalities of applications. The classifiers are another component which is the xml files classified in face classifiers and emotion classifiers. Another box is the Android package on which the application will be deployed and then will be provided to client for using it frequently.



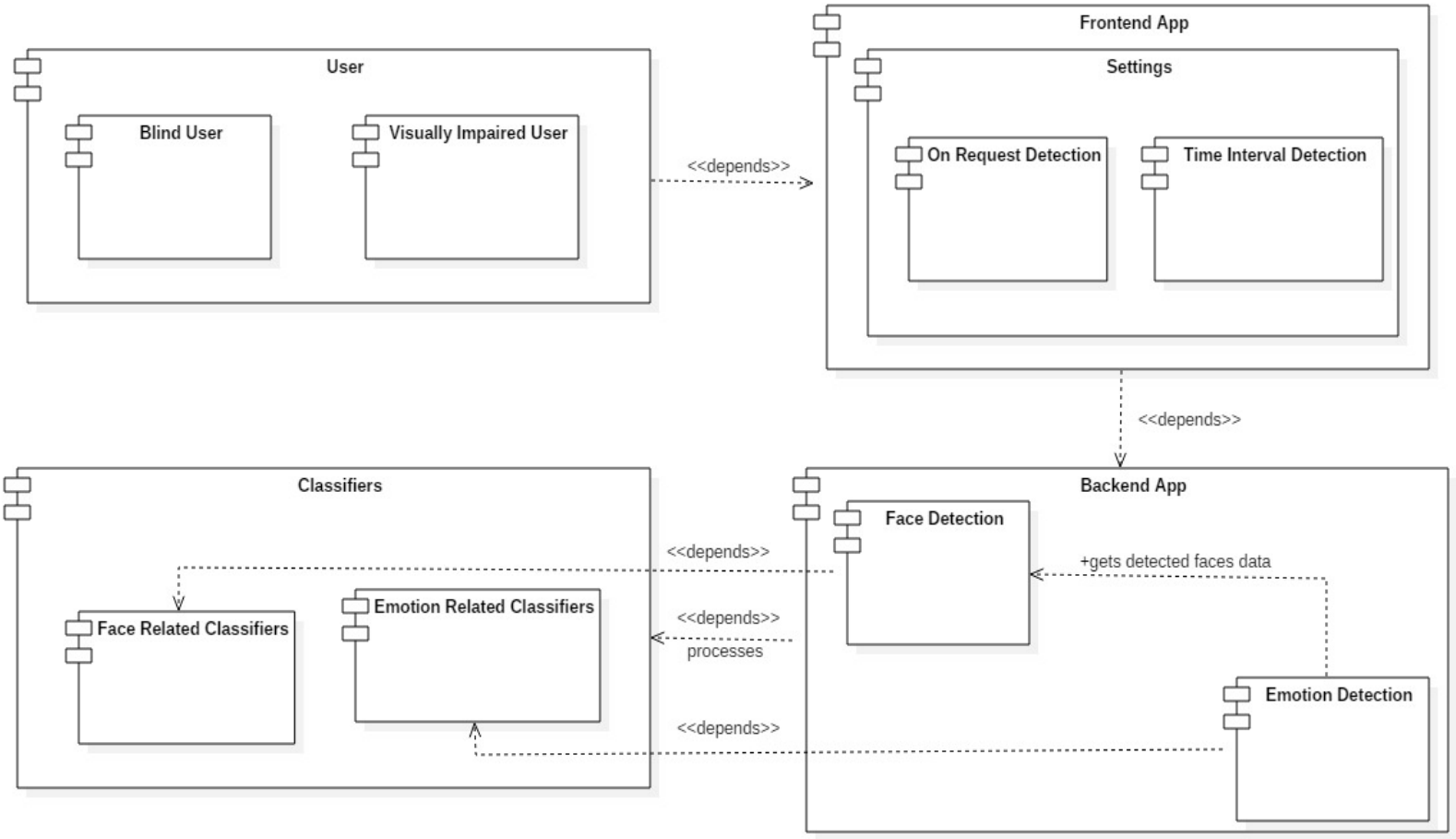
4.1.2.3 Component Diagram

Component diagram is an exceptional sort of diagram in universal modeling language. The reason for existing is additionally unique in relation to every single other diagram talked about in this way. It doesn't depict the usefulness of the framework however it portrays the segments used to make those functionalities. There are number of components in this application. One of the components is the Frontend app component which encapsulates all the appearance and interface related terms and functionalities. Second component is the Backend App component which encapsulates all the hidden functionalities and processing work. The third component is the classifiers which can be further divided into subcomponents which are explained in the next section. The classifier component helps the Backend App component to detect faces and emotions. The User component specifies the respective users of the EyeHope application.



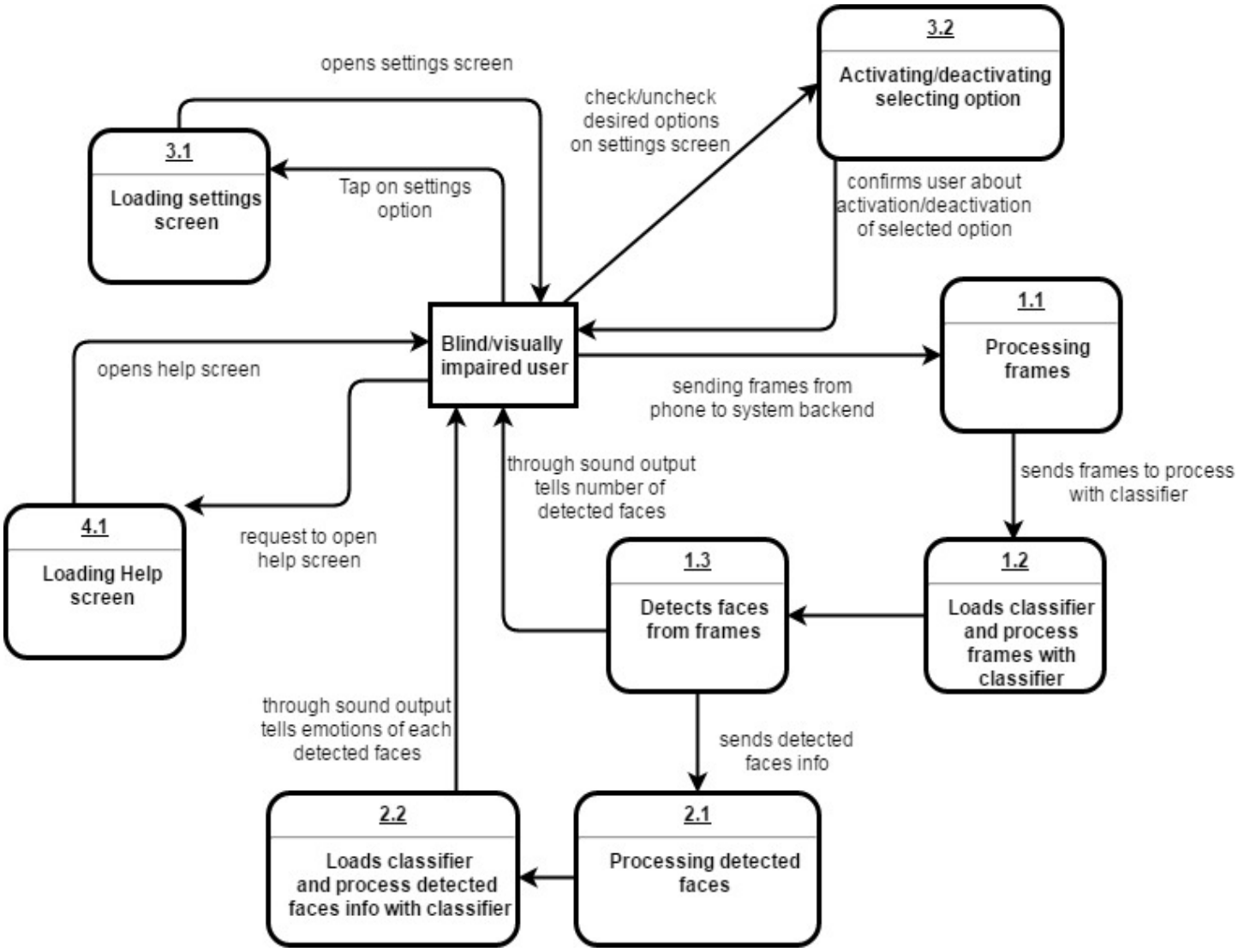
4.1.2.4 Sub-Component Level Architecture

There are number of components in this application. The Frontend app component, which encapsulates all the appearance and interface related terms and functionalities, can be decompose into sub-components such as settings and in settings also sub classifies into on request detection and time interval detection. Second component is the Backend App component, which encapsulates all the hidden functionalities and processing work, can be decompose into sub-components such as Face detection component and Emotion detection component. The third component is the classifier which can be further divided into subcomponents such as Face related classifier component and Emotion related classifier component. The cascade classifiers can be viewed as an object specific focus-of-attention mechanism which unlike previous approaches provides statistical guarantees that discarded regions are unlikely to contain the object of interest [5]. The classifier component helps the Backend App component to detect faces and emotions. The User component contains specifies the respective users of the EyeHope application.



4.1.3 Data Flow Diagram (DFD)

The data flow diagram below shows a deep level understanding about flowing of data among different processes. The data flow between processes 1.1, 1.2 and 1.3 shows successful completion of face detection process of the system. Whereas, data flow between processes 1.3, 2.1 and 2.2 shows successful completion of emotion detection process of the system. Data flowing through process 4.1 shows the procedure for opening help screen whereas data flowing through processes 3.1 and 3.2 shows how the settings screen will be open user as well as how the data will flow while changing and activating the changed settings of the system.

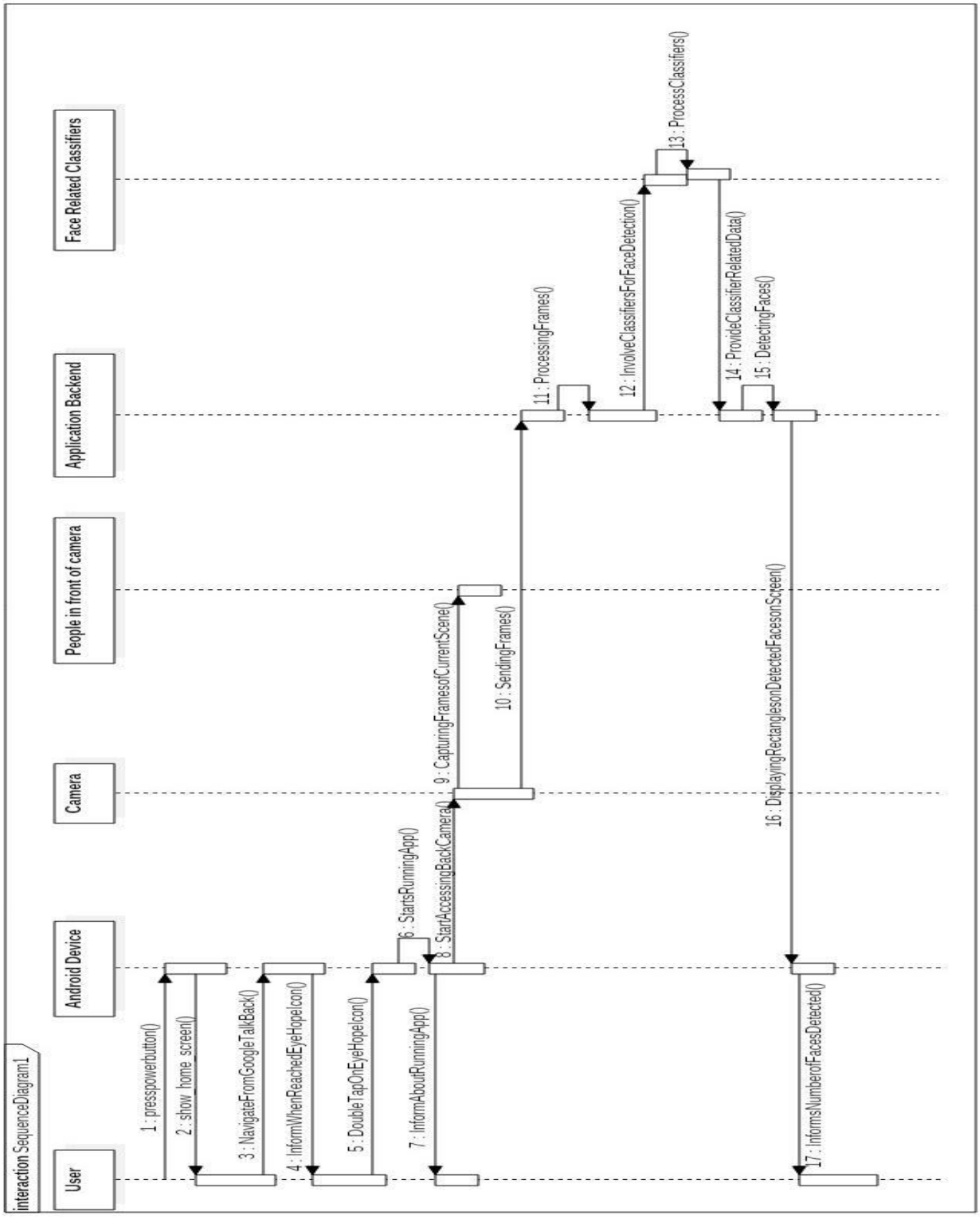


4.1.4 Sequence Diagram

The sequence diagram of EyeHope contains a detailed sequence of actions which are closely related to each other. The sequence diagram of EyeHope Application is divided into multiple diagrams according to the feature it provides, which are shown below,

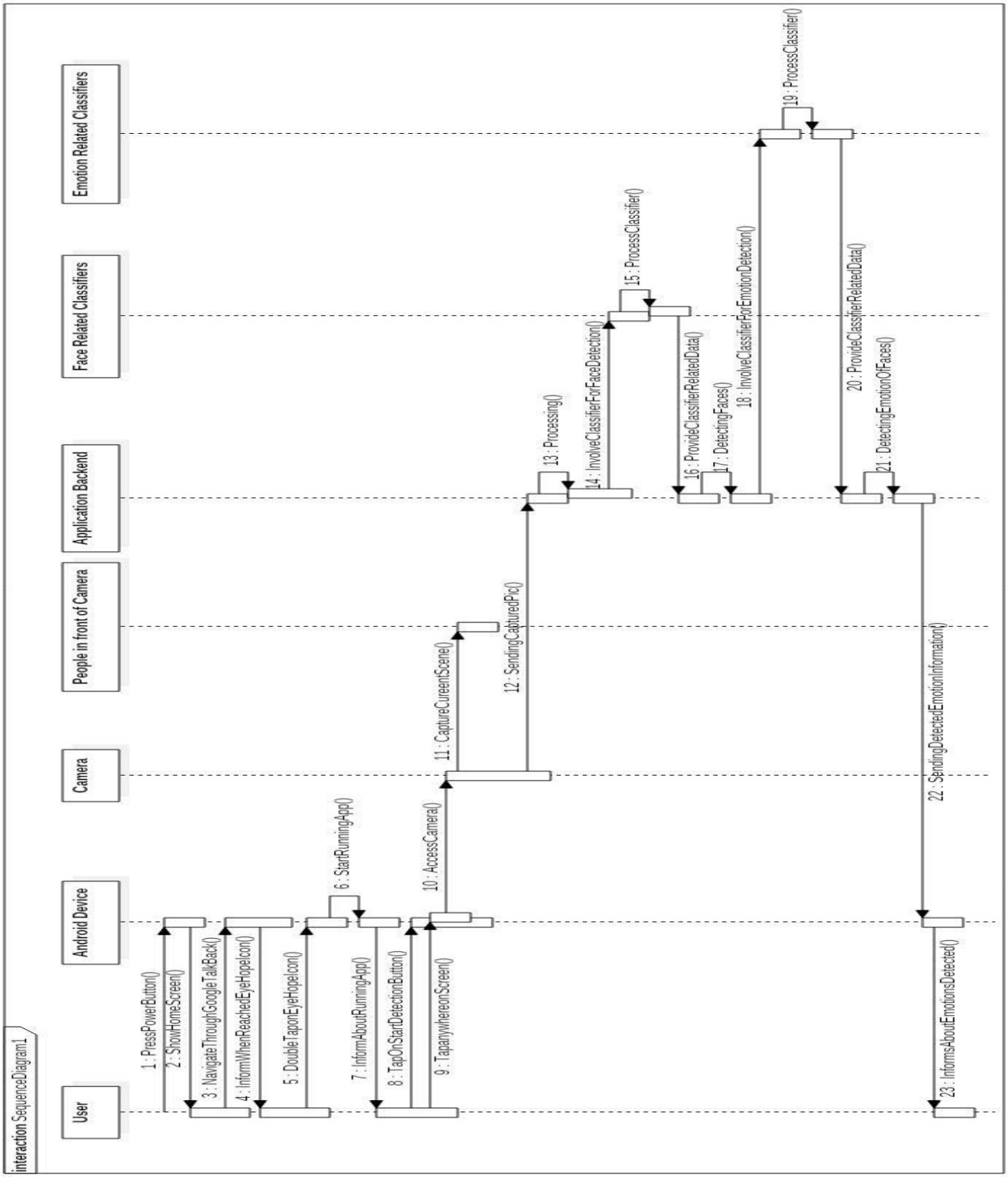
1) Sequence Diagram for Face Detection

One of the important features of EyeHope application is to detect the number of people which are present in front of camera. From this the user will be acknowledged about how many people are in front of user. The sequence diagram for face detection includes the actions performed by Eyehope application to successfully detect the faces and inform the user about the number of faces detected. In the given diagram, the flow is started from opening the android device from scratch. The user will navigate towards the EyeHope application icon through GoogleTalkBack accessibility mode. The User taps the icon to run the application. When the user taps anywhere on screen the frames are captured by camera and are forwarded to backend of the application where the classifiers are accessed and face related information from face classifier and processed. From this processing the backend application will compute the number of faces detected and will provide the output to user through sound. The detailed sequence diagram for face detection is shown below,


interaction SequenceDiagram1
User
Android Device
Camera
People in front of camera
Application Backend
Face Related Classifiers
1 : presspowerbutton()
2 : show_home_screen()
3 : NavigateFromGoogleTalkBack()
4 : InformWhenReachedEyeHopeIcon()
5 : DoubleTapOnEyeHopeIcon()
6 : StartsRunningApp()
7 : InformAboutRunningApp()
8 : StartAccessingBackCamera()
9 : CapturingFramesofCurrentScene()
10 : SendingFrames()
11 : ProcessingFrames()
12 : InvolveClassifiersForFaceDetection()
13 : ProcessClassifiers()
14 : ProvideClassifierRelatedData()
15 : DetectingFaces()
16 : DisplayingRectanglesonDetectedFacesonScreen()
17 : InformsNumberofFacesDetected()

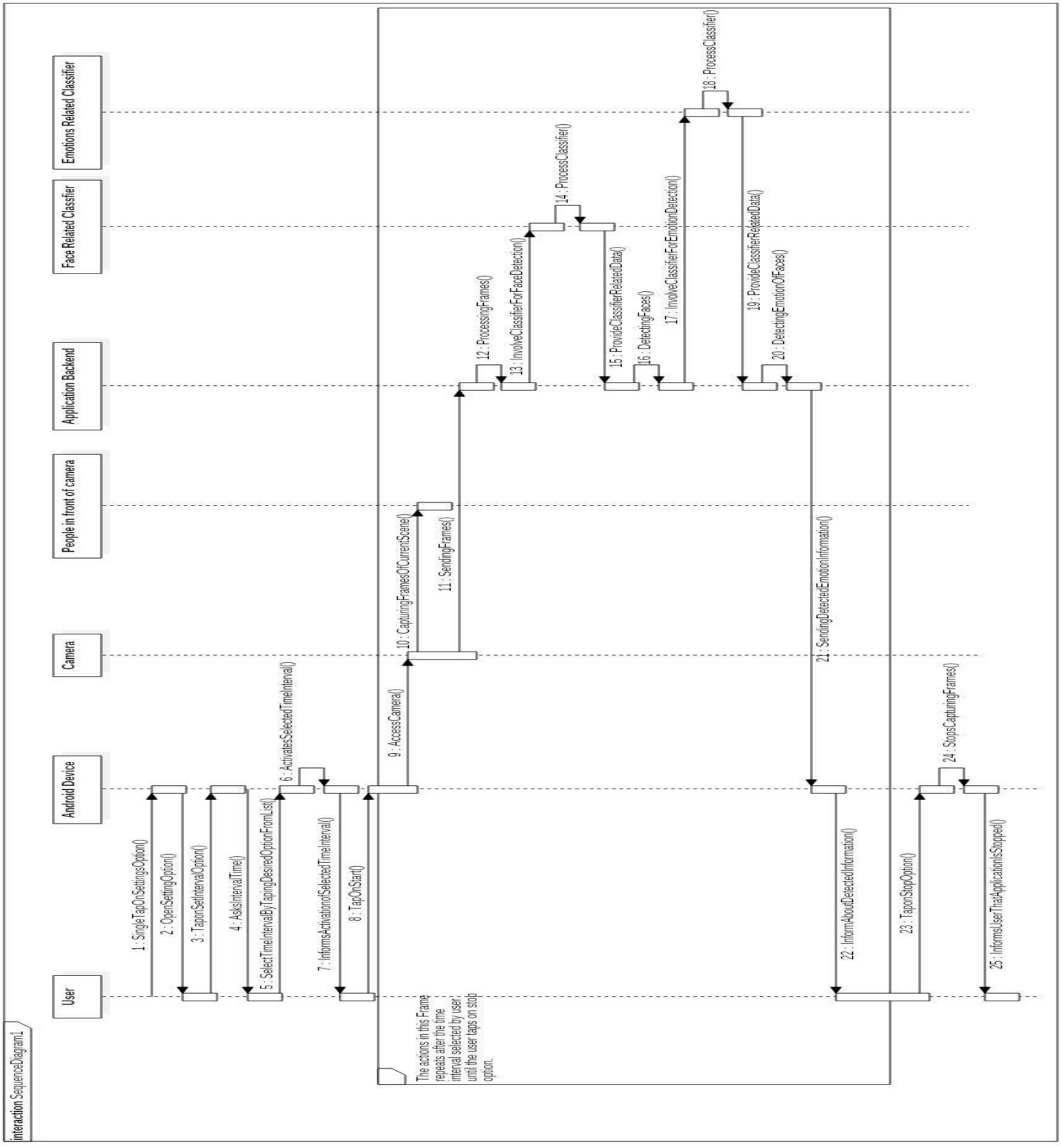
2) Sequence Diagram for Emotion Detection through On-request Option

For getting information of detected faces and emotions of detected faces can be requested by users through two options. One of those two options is On request option in which, when the user randomly taps on any part of the screen, the frames (on tap) are captured and are forwarded to backend of the application where the classifiers are accessed and face related information from face classifier are processed. From this processing, the backend application will detect the faces from captured frames and will process these detected faces information with emotion related classifier to detect the respective emotion of the detected faces. The detailed sequence diagram for emotion detection through on-request option is shown on next page,


interaction SequenceDiagram1
User
Android Device
Camera
People in front of Camera
Application Backend
Face Related Classifiers
Emotion Related Classifiers
1 : PressPowerButton()
2 : ShowHomeScreen()
3 : NavigateThroughGoogleTalkBack()
4 : InformWhenReachedEyeHopeIcon()
5 : DoubleTaponEyeHopeIcon()
6 : StartRunningApp()
7 : InformAboutRunningApp()
8 : TapOnStartDetectionButton()
9 : TapanywhereonScreen()
10 : AccessCamera()
11 : CaptureCureentScene()
12 : SendingCapturedPic()
13 : Processing()
14 : InvolveClassifierForFaceDetection()
15 : ProcessClassifier()
16 : ProvideClassifierRelatedData()
17 : DetectingFaces()
18 : InvolveClassifierForEmotionDetection()
19 : ProcessClassifier()
20 : ProvideClassifierRelatedData()
21 : DetectingEmotionOfFaces()
22 : SendingDetectedEmotionInformation()
23 : InformsAboutEmotionsDetected()

3) Sequence Diagram for Emotion Detection through Time Interval Option

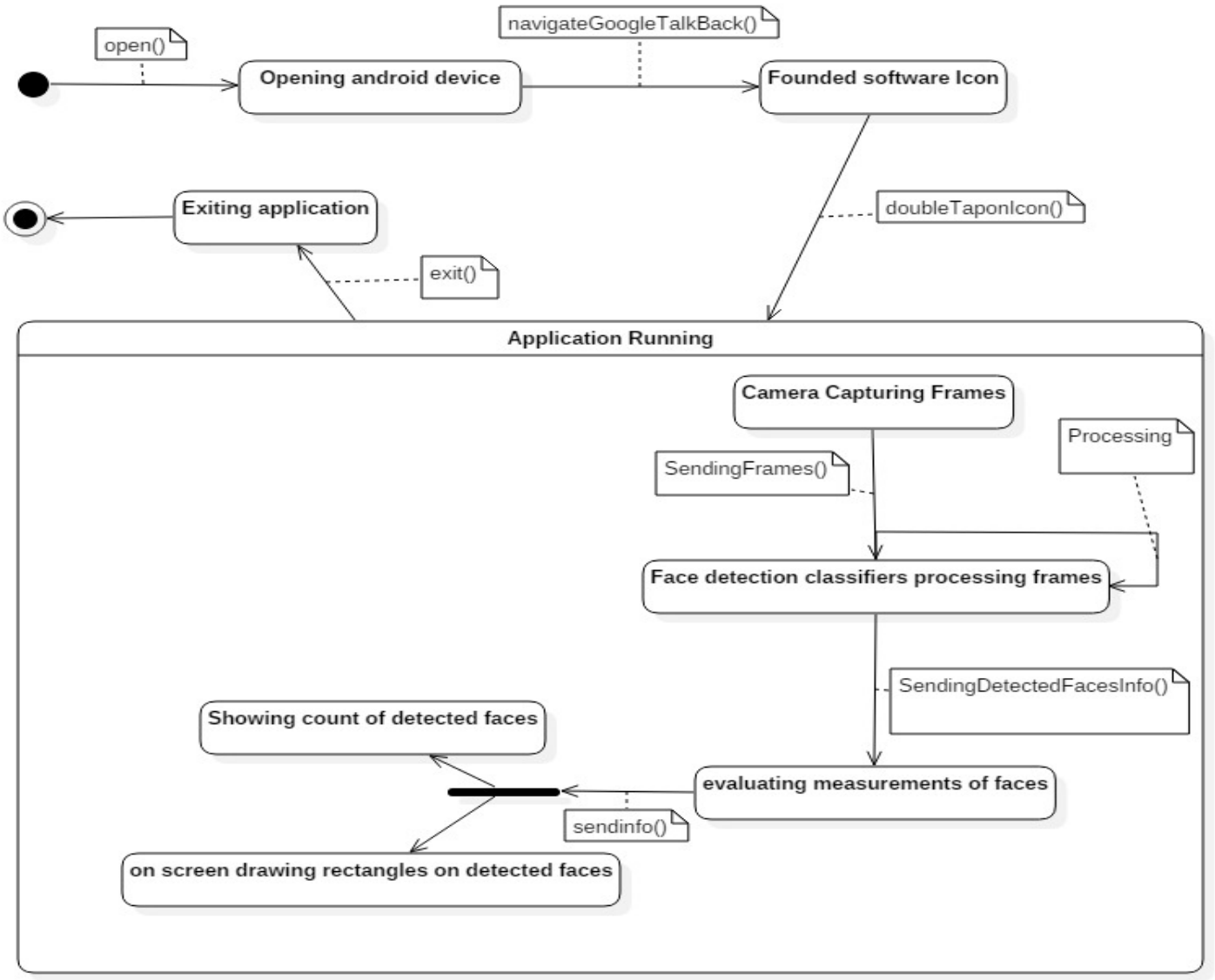
For getting information of detected faces and emotions of detected faces can be requested by users through two options. One of those two options is set interval option. The user accesses the settings option and selects set interval option and selects the desired time interval option (should be greater than equal to five). After activating the time interval option, between this time intervals, the frames are captured and are forwarded to backend of the application where the classifiers are accessed and face related information from face classifier are processed. From this processing, the backend application will detect the faces from captured frames and will process these detected faces information with emotion related classifier to detect the respective emotion of the detected faces. This whole cycle will continue in for the given time interval until the current state of application is turned off by tap on back button. The detailed sequence diagram for emotion detection through time interval option is shown on next page.


interaction SequenceDiagram1
User
Android Device
Camera
People in front of camera
Application Backend
Face Related Classfier
Emotions Related Classifier
1 : SingleTapOnSettingsOption()
2 : OpenSettingOption()
3 : TaponSetIntervalOption()
4 : AskIsIntervalTime()
5 : SelectTimeIntervalByTapingDesiredOptionFromList()
6 : ActivatesSelectedTimeInterval()
7 : InformsActivationOfSelectedTimeInterval()
8 : TapOnStart()
9 : AccessCamera()
10 : CapturingFramesOfCurrentScene()
11 : SendingFrames()
12 : ProcessingFrames()
13 : InvolveClassifierForFaceDetection()
14 : ProcessClassifier()
15 : ProvideClassifierRelatedData()
16 : DetectingFaces()
17 : InvolveClassifierForEmotionDetection()
18 : ProcessClassifier()
19 : ProvideClassifierRelatedData()
20 : DetectingEmotionOfFaces()
21 : SendingDetectedEmotionInformation()
22 : InformAbooutDetectedInformation()
23 : TaponStopOption()
24 : StopsCapturingFrames()
25 : InformsUserThatApplicationIsStopped()
The actions in this Frame repeats after the time interval selected by user until the user taps on stop option.

4.1.5 State Chart Diagram

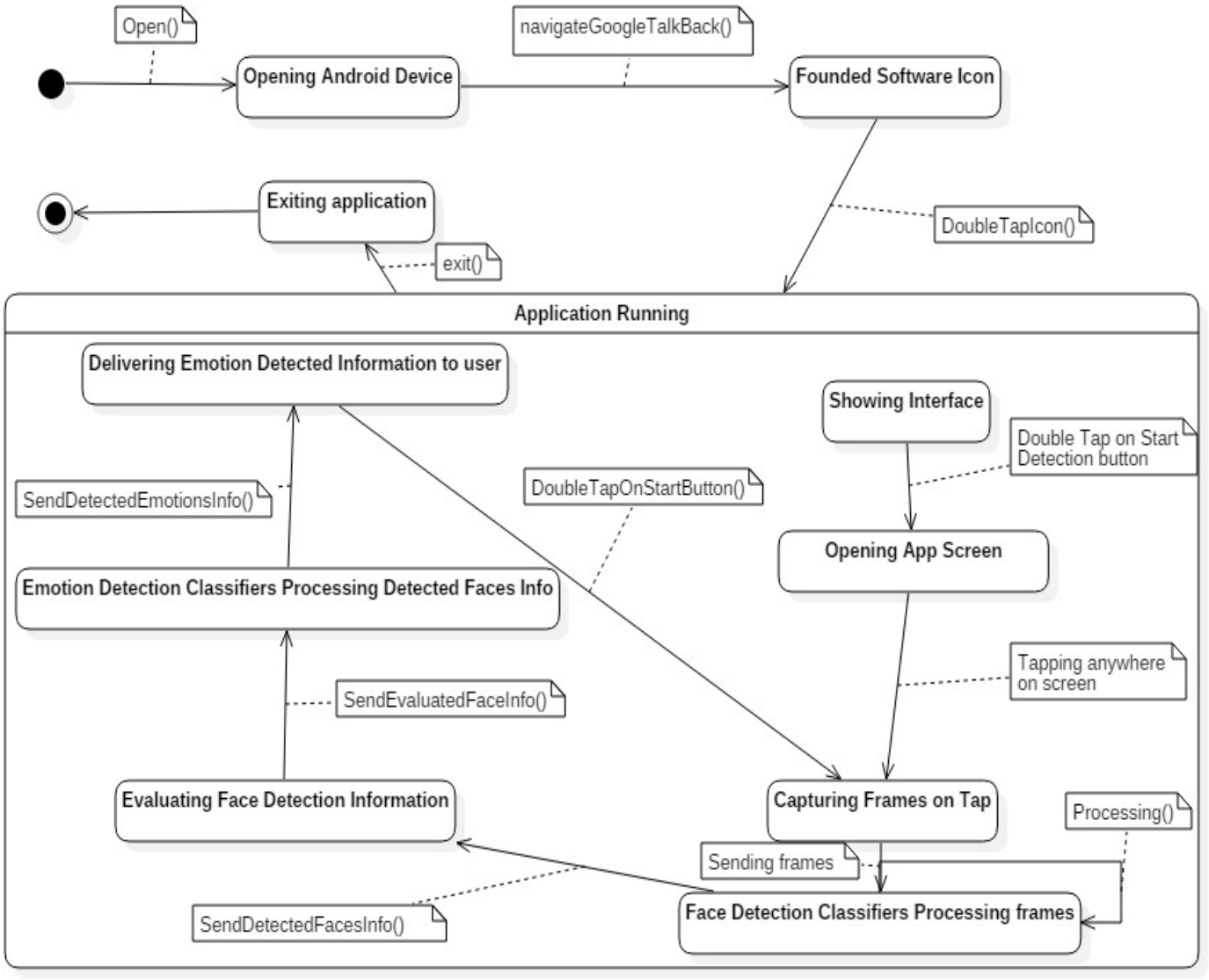
1) State Chart Diagram for Face Detection

The state chart diagrams are those which reflect the flow of actions and states, which are caused after applying the actions, professionally. The flow of actions and states starts from opening the android device from scratch. The user will navigate towards the EyeHope application icon through GoogleTalkBack accessibility mode; from this the user will reach the EyeHope application icon. The User taps the icon to run the application. When the user taps anywhere on screen the frames are captured by camera and are forwarded to backend of the application where the classifiers are accessed and face related information from face classifier and processed. From this processing state, the backend application will compute the number of faces detected and will provide the output to user through sound. The state chart diagram for face detection is shown on next page.



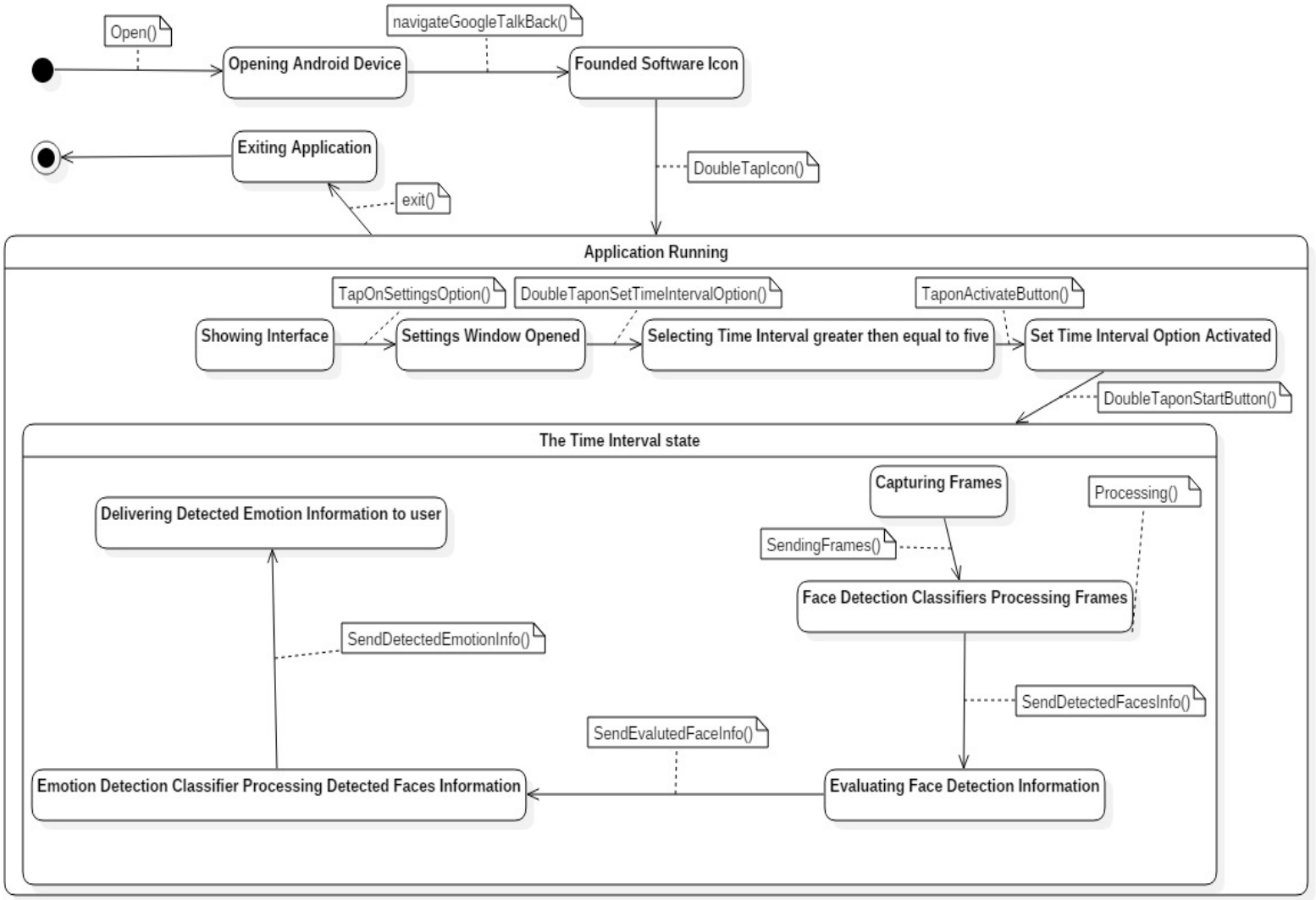
2) State Diagram for Emotion Detection through On-request Option

For getting information of detected faces and emotions of detected faces can be requested by users through two options. One of those two options is On request option in which, when the user randomly taps on any part of the screen (action), the frames (on tap) are captured (state) and are forwarded (action) to backend of the application where the classifiers are accessed (state) and face related information from face classifier are processed (state). From this processing (state), the backend application will detect the faces (action) from captured frames and will process these detected faces information with emotion related classifier to detect the respective emotion of the detected faces. The state chart diagram for emotion detection through Onrequest option is shown on next page.



3) State Chart Diagram For Emotion Detection Through Time Interval Option

For getting information of detected faces and emotions of detected faces can be requested by users through two options. One of those two options is set interval option. The user accesses the settings option (state) and selects set interval option (action) and selects the desired time interval option (action) (should be greater than equal to five). After activating the time interval option (state), between this time intervals, the frames are captured and are forwarded to backend of the application (action) where the classifiers are accessed (state) and face related information from face classifier are processed (state). From this processing, the backend application will detect the faces from captured frames (state) and will process these detected faces information with emotion related classifier (state) to detect the respective emotion of the detected faces (action). This whole cycle will continue in for the given time interval until the current state of application is turned off by tapping on back button. The state chart diagram for emotion detection through time interval option is shown on next page.



5. SOFTWARE TESTING

5.1 Test Plan Strategy

Software testing plan is utilized to guarantee that every one of the parts and modules of the product are working consummately without blunders. Giving conceivable contributions to the framework, the yields are checked in the event that they are as what was relied upon to be the errand of the framework.

As our testing strategy, we used Unit testing to individually test all the separately developed modules such as face detection/ detecting number of people module, emotion detection module, Help module and settings module. As a second step, the Integration testing was carried out when the modules were being integrated with each other. As a third and final step, when the integration testing was finally completed, the System testing was carried out to ensure whether the whole system is working in a proper flow or not.

5.2 Unit Testing

Definition:
Unit Test consists of testing individual programs or subroutines as they are written instead of testing the entire system after it has been written. The testing of the smaller building blocks is done first and then these blocks are combined and tested. Unit testing means testing each function independently to verify correct processing in a stand-alone environment.

Participants*:*
Aamir Amin Khimani, Mahzain Malik

Methodology:

- Each and every module was separately tested right after it was developed.
- When testing was being held on a particular module, all of its functionalities were separately tested by temporarily disabling the others and focusing on the testing of only that particular module.
- In face detection/ detecting number of people module, different types of images and video streaming were used to check whether it is detecting each face in the image/videos clearly or not.
- In emotion detection module, the emotions of each detected faces were checked whether they are correctly detected or not.
- In help module, it was checked whether the text on the screen is somehow visible to visually impaired users and whether TalkBack mode properly dictates the text which is written on the screen.
- In settings module, it was checked whether the ticked/un-ticked options were properly activated/deactivated in the system.

5.3 Integration Testing

Definition:
Integration Testing is a level of software testing where individual units are combined and tested as a group. The purpose of this level of testing is to expose faults in the interaction between integrated units. Test drivers and test stubs are used to assist in Integration Testing.

Participants:
Hammad Mubarak, Saman Karim

Methodology:

- While integrating any of the two modules, it was tested that there are no such cases in which one module disrupts or interrupts the behavior of other module.
- The back-end module (Processing) and front-end module (Application) were properly tested for the synchronization.
- It was also tested that transferring of information between two modules is consistent and there is no exceptional or abnormal behavior.

5.4 System Testing

Definition:
System Testing is a level of the software testing where complete and integrated software is tested. The purpose of this test is to evaluate the system's compliance with the specified requirements.

Participants:
Hammad Mubarak, Aamir Amin Khimani, Mahzain Malik, Saman Karim

Methodology:

- The complete system was tested, with all the operations being carried out in order that the whole system should work properly.
- Different types of images and videos were provided as input and on these inputs, it was checked and evaluated whether the output is same as expected, after processing the input throughout the system.

The real time frames were being fetched by the mobile camera and were processed by backend of the application. The application first detects the faces in the frame and then detects emotions of detected faces. In system testing, the process stated above is properly been checked so that the system should become errorless.

6. SYSTEM USER GUIDE

6.1 List of Interaction Modules

The interaction modules for the EyeHope system are as follows,

1) Main application screen containing three options

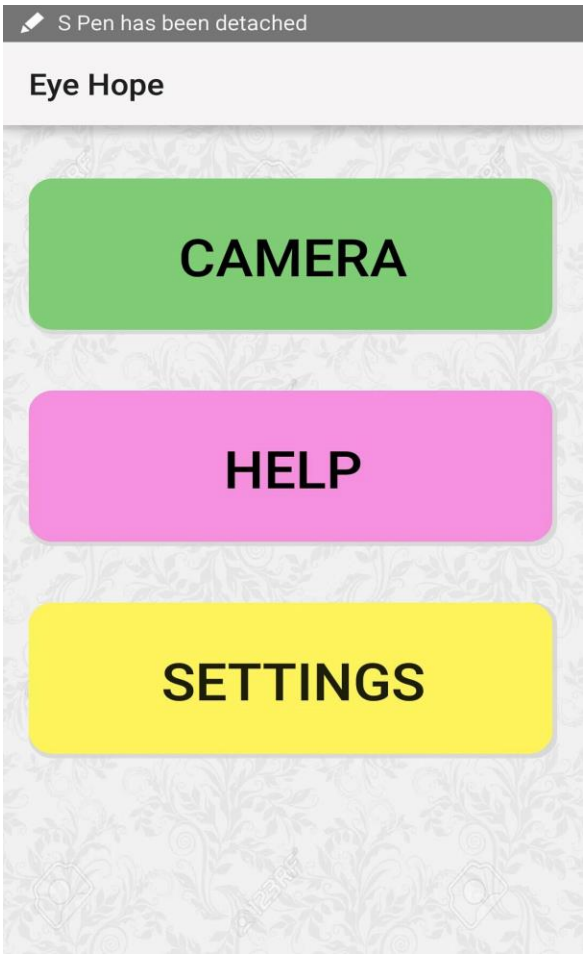
2) Camera screen where faces and emotions will be detected

3) Settings screen where filtering of particular emotions and detections (without interval and with interval) are done.

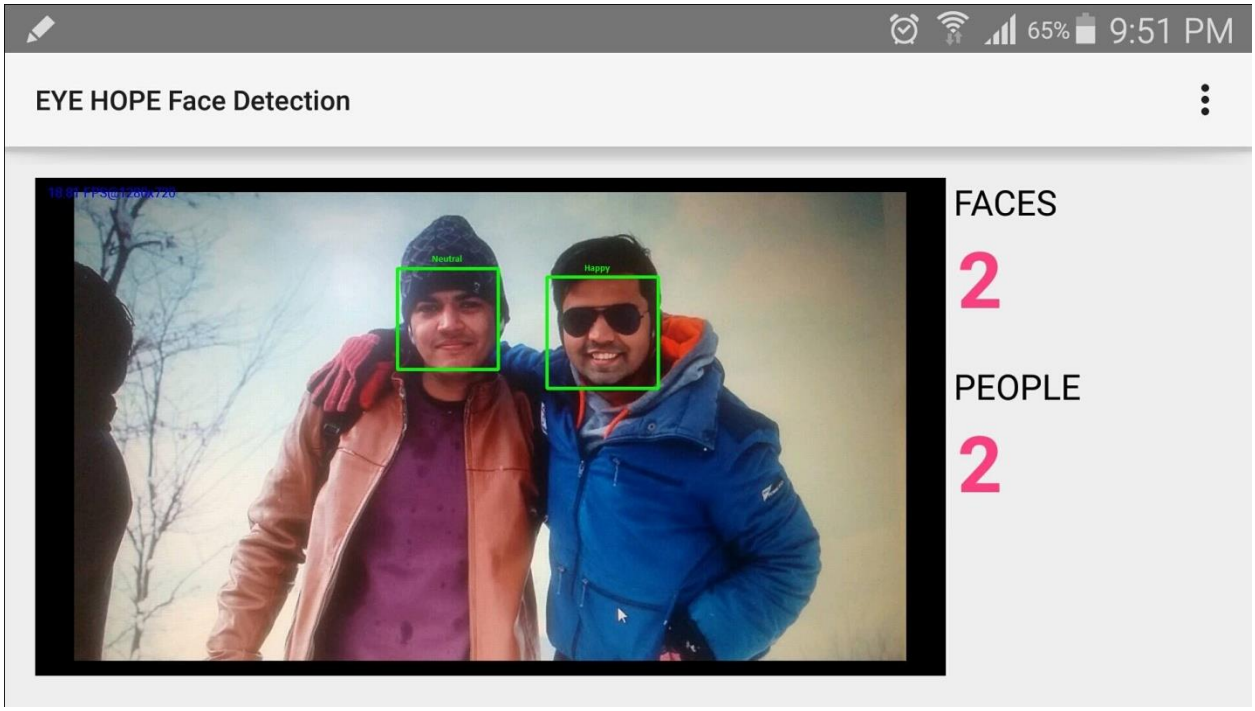
4) Help screen where user can be guided how the system works.

6.2 Snapshots of Interaction Modules

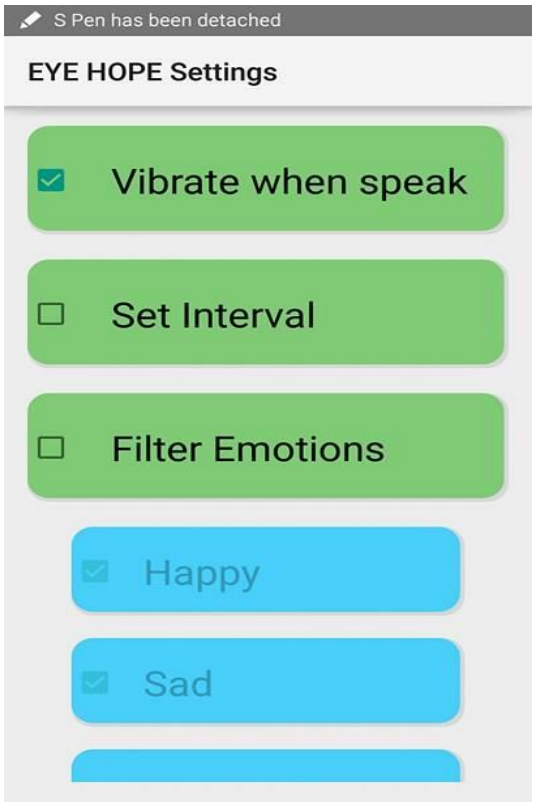
1) Home Screen Module



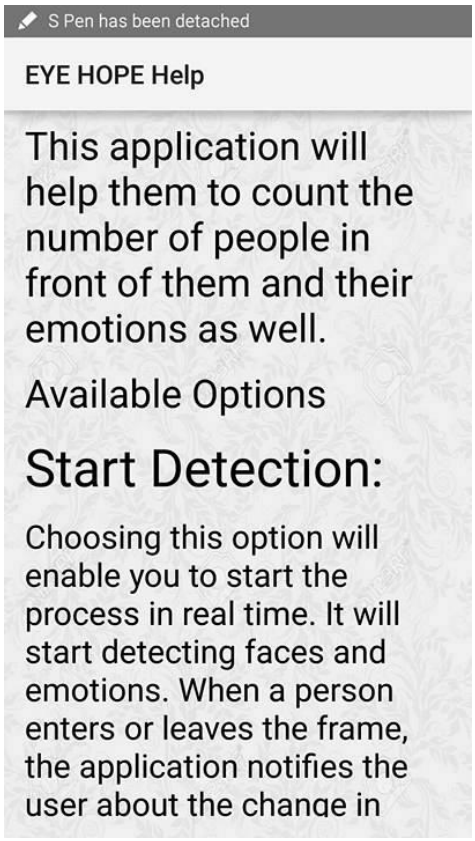
2) Camera Module

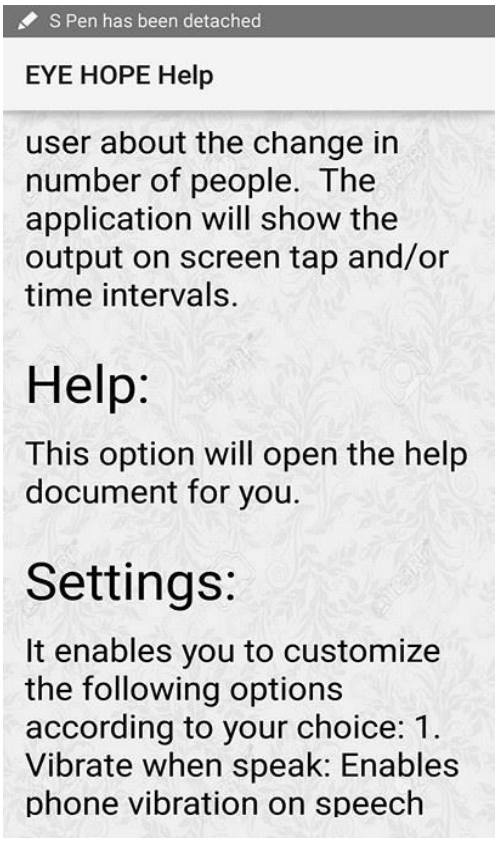


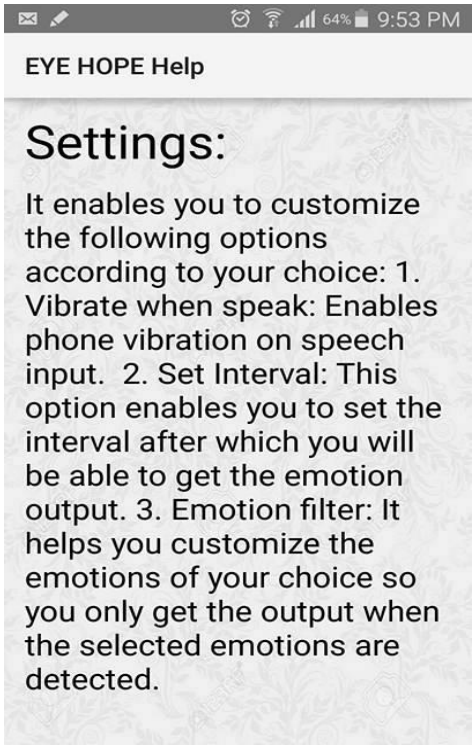
3) Settings Module



4) Help Module







6.3 Inputs and Expected Outputs

Module 1: Main Application screen

1) Inputs

a. User double taps on the application launch icon

b. User double taps on Camera option

c. User double taps on Settings option

d. User double taps on Help option

2) Means of Inputs

The inputs specified above are given by user through touching (double taping) the particular areas of screen (each particular option is in the particular area of screen).

3) Outputs

a. Main application screen will appear on the mobile screen.

b. A screen appears on which much of the left side area shows camera results and fetches frames, right side area shows number of people detected.

c. Opens a settings screen where it will show filtering options.

d. Opens a Help screen.

Module 2: Camera Screen

1) Inputs

a. User double taps on three vertical dots options on top right side.
b. User keeps the phone in front of their own face (facing the screen side) in horizontal direction.

2) Means of Inputs

a. This input is given by user through touching (double taping) the particular areas of screen.

3) Outputs

a. Opens a small dialogue box which requires frame size (face size) to be detected.

b. Voice output for number of faces detected.

Module 3: Settings Screen

1) Inputs

(User double taps on settings screen to tick/un-tick the given options)

a. User double taps on 'Vibrate when speak' to tick/un-tick this option.

b. User double taps on 'Set Interval' to tick/un-tick this option.

c. User double taps on 'Filter Emotions' to tick/un-tick this option and double taps on different desired options to tick or un-tick these options.

2) Means of Inputs

The inputs specified in this module are given by user through touching (double taping) the particular areas of screen (each particular option is in the particular area of screen).

3) Outputs

a. 'Vibrate when speak' option activates, as when output for numbers of detected faces and its emotions generates the device will also vibrate.

b. 'Set Interval' option activates, as the voice output generates after every particular time interval.

c. The emotions selected by user are activated and the output will consist of only those selected emotions.

Module 4: Help Screen

1) Inputs

a. User does a single tap on screen.

2) Means of Inputs

For this module, a user gives input through touching (single taping) the screen.

3) Outputs

a. TalkBack reads the user guide written on the screen.

7. CRITICAL EVALUATION

7.1 Success Criteria

Our research was mainly focused on classification of 6 universal facial expressions of any human. We have spent extensive time improvising our techniques and analyzing the accuracy. The factors included in the success criteria for EyeHope application are as follows,

1) Maximum accuracy for detecting faces and emotions.

2) Number of downloads.

3) Deliverables to be submitted before deadline.

4) Feedback from targeted users.

7.2 Degree of Success

Initially, we have achieved an accuracy of 67% on a data set interpreted by humans. Handling the noises in dataset's and over-fitting of different classifiers our accuracy was close to 73.7%. To test the credibility of the system and to avoid any discrepancy in it, we have also tested our application on the dataset's collected and designed by ourselves. After many optimizations and observing our application by applying various approaches, our final accuracy was around 80.53%. This reflects that the emotions detected by EyeHope application is are more likely to match the interpretation of a normal human observer.

7.3 Assumptions and Limitations

It is assumed that the accessibility feature (talk back) is turned on while users use this application. TalkBack is an availability administration that blinds and vision-debilitated clients associate with their gadgets. It is an interesting accessibility service that helps blinds and visually impaired people to use and communicate with their device properly and conveniently. TalkBack includes talked feedback, hearing feedback, and vibration feedback to your gadget. The Google talk back is compatible with almost every android device. The face detection feature is only applicable for frontal faces which mean that while detecting the faces on real time, the system will detect only those faces which are exactly in front of camera frame and are exactly facing the smart phone camera.

7.4 Resources

1) Hardware Resources

a. A Smartphone with embedded camera to capture frames.

b. Pair of Ear phones to facilitate the blind/visually impaired users with the output in speech (optional).

2) Software Resources

a. Smartphone supporting android versions 4.1 (Jelly Bean) and above.

b. OpenCV for Image processing.

c. Haar Cascade files for face detection.

3) Dataset Resources

a. Michael J. Lyons, Shigeru Akemastu, Miyuki Kamachi, Jiro Gyoba. Coding Facial Expressions with Gabor Wavelets, 3rd IEEE International Conference on Automatic Face and Gesture Recognition, pp. 200-205 (1998).

b. Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00), Grenoble, France, 46-53.

7.5 Future Enhancements

The system of EyeHope application can be enhanced by integrating more features in it such as,

1) Text detection by processing the frames having any kind of text in it.

2) Describing the scenario in the frame.

3) Face Recognition.

4) Object Detection.

5) Object Recognition.

The system can also be enhanced by modifying or polishing the current features like,

1) The frontal face detection feature can be extended by detecting the face even if camera is facing only left side of the face or only right side of the face.

2) The settings module can also be modified to implement more filtering options.

8. REFERENCES

1. Monica Chillaron et al.Face detection and recognition application for Android". In: (2015). doi: https://www.researchgate.net/publication/284253914 _ Face _ detection _ and _ recognition _ application _ for _Android.
2. H. Bourlard D. Chen J-M. Odobez. Text detection and recognition in images and video frames," in: 37 (2004), pp. 595-608.
3. Sh. Krishna et al. A wearable face recognition system for individuals with visual impairments," in: (October 2005), pp. 106-113.
4. P. Viola and M. Jones. Robust Real-Time Face Detection," in: 57 (2004), pp. 137-154.
5. A. L. Yuille X. Chen. A Time-Effect Cascade for Real-Time Object Detection: With applications for the visually impaired," in: (June 20, 2005).
6. Emotion Recognition: A Pattern Analysis Approach by Amit Konar, Aruna Chakraborty
7. DEVELOPING AN IMAGE RECOGNITION ALGORITHM FOR FACIAL AND DIGIT IDENTIFICATION Christian Cosgrove, Kelly Li, Rebecca Lin, Shree Nadkarni, Samanvit Vijapur, Priscilla Wong,Yanun Yang, Kate Yuan, Daniel Zheng
8. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," NIPS, vol. 1, p. 4, 2012.
9. A. Yao, J. Shao, N. Ma, and Y. Chen, "Capturing au-aware facial features and their latent relations for emotion recognition in the wild," in Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, ICMI '15, (New York, NY, USA), pp. 451–458, ACM, 2015.
10. C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," Image and Vision Computing, vol. 27, no. 6, pp. 803 – 816, 2009.
11. S. Kahou, V. Michalski, K. Konda, R. Memisevic, and C. Pal, "Recurrent neural networks for emotion recognition in video," ICMI, pp. 467–474, 2015.
12. Z. Yu and C. Zhang, "Image based static facial expression recognition with multiple deep network learning," in Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, ICMI '15, (New York, NY, USA), pp. 435–442, ACM, 2015.
13. B. Kim, J. Roh, S. Dong, and S. Lee, "Hierarchical committee of deep convolutional neural networks for robust facial expression recognition," Journal on Multimodal User Interfaces, pp. 1–17, 2016.

14. Confusion Matrix-based Feature Selection by Sofia Visa, Brian Ramsay, Anca Ralescu, Esther van der Knaap

15. A Review of the Literature on Emotional FacialExpression and Its NatureP›nar Dursun*, Murat Emül**, Faruk Gençöz***

16. A Literature review on Emotion Recognition System Using Various Facial Expression Dipika Raval1, Mukesh Sakle2

17. S Sukanya Sagarika ;Pallavi Maben, Laser Face Recognition and Facial Expression Identification using PCA, 2014 IEEE

18. S L Happy;Anjith George;Aurobinda Routray, "A Real Time Facial Expression Classification System Using Local Binary Patterns.," 2012 IEEE

19. Aruna Bhadu, Rajbala Tokas, Dr. Vijay Kumar, "Facial Expression Recognition Using DCT, Gabor and Wavelet Feature Extraction Techniques," International Journal of Engineering and Innovative Technology, Volume 2, Issue 1, July 2012.

20. A Literature Review on Facial and Expression Recognition with Advanced Image Descriptor Template by Pranjali S. Bele and Prakash S. Mohod

21. Real-time Mobile Facial Expression Recognition System – A Case Study Myunghoon Suk and Balakrishnan Prabhakaran Department of Computer Engineering The University of Texas at Dallas, Richardson, TX 75080

22. P. Ekman and W. Friesen. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto, 1978.

23. https://karanjthakkar.wordpress.com/2012/11/21/what-is-opencv-opencv-vs-matlab/

24. https://www.researchgate.net/post/Which_is_the_best_opencv_or_matlab_for_image_processing