Software Requirement Specification Document for Gestify

Seif Hisham, Ahmed Amr, Marina Maurice, Kirolos Emad
Supervised by: Dr. Salwa Osama, Eng. Donia Mohamed

April 30, 2024

Table 1: Document version history

Version	Date	Reason for Change
1.0	6-Jan-2024	SRS First version's specifications are defined.
1.1	9-Jan-2024	Class Diagram Created
1.2	11–Jan-2024	Context diagram, use case modifications, overview diagram description
1.3	14–Jan-2024	Operational Scenarios- Detailed Functional requirements

GitHub: TouchPoint-Detection-Advancements-in-Projector-Camera-HCI

Contents

Abstract

The main idea of this project is to study the contactless control over a PC connected to any projector using only hand gestures. Moreover, reducing the cost associated with smart boards and many other solutions like touch projectors, which are still expensive compared to our software that enables anyone with his smartphone's camera, a projector, and a simple setup to have an interactive surface anywhere controlled remotely with just hand gestures, in addition to face recognition feature - which uses YOLO algorithm - that enables gesture customization makes it the best affordable solution in the market that can be used for teaching and presentations.

1 Introduction

1.1 Purpose of this document

The purpose of this document is to specify and explain our software implementation, algorithms, diagrams for illustration, and finally, mention our approach. Gestify software aims to provide a cost-effective solution for more interactive presentations without any hardware interference for instance the mouse; for clicking or dragging and dropping any icon/folder anywhere. Also, there is a feature where the user can use his hand gestures to give commands to his laptop(turn up/down or mute PC sound) while presenting, making it a more interactive, easy-to-use, and cheaper solution accessible for everyone.

1.2 Scope of this document

This paper examines similar systems to the Gestify system, in addition, it illustrates the overview, scope, and context of system design and addresses the objectives and features offered by the software. This document also goes through Gestify's functional and non-functional requirements, design limitations, data design, and the software's class diagram in detail. Finally, this paper presents the software's timeline.

1.3 Business Context

In the last decade, Touch projectors have become more popular, especially in schools and businesses. Nowadays people want more interactive displays for teaching and presentations more than before, and touch projectors are meeting that demand. To make this technology accessible to many, we're focusing on making touch projectors cost-effective. This means they are not too expensive, making them a good choice for schools and businesses that want interactive learning tools without spending too much money. [1]

2 Similar Systems

2.1 Academic

1- Making any planar surface into a touch-sensitive display by a mere projector and camera[2]:

The objective of the accompanying paper is to change over any tabletop into a touch-delicate screen when the projection is enlightened, where the framework is made out of a video projector and camera. It additionally handles a few issues related with the arrangement they developed. For example: The liquid-crystal display (LCD) panel, which makes the device less portable, was mentioned in recent research papers as working on finger tracking but ignoring finger clicking detection algorithms. Later research [3], [4] attempted to detect finger clicking, but the user had to wait a while for his click to be detected. For the investigation, they utilized a DLP projector with a local goal of 640×480 and a point of interaction for firmware design (TI DLP Pico Projector Improvement Pack 2), notwithstanding a camera of 648×488 goal (Point Dark FL3-FW-03S1C camera with Myutron FV0622 f6mm focal point). They utilized C++ programming language involving Intel OpenCV Library for handling time assessment, the general time taken for hand division in 2D picture, fingertip restriction, and contact discovery are under 20ms. The framework was designed for a functioning distance of around 500mm, making a 15-inch projection region, they referenced likewise in the event that a short-toss projector was utilized a greater projection region could be procured.

2- Any-Wall Touch Control System With Switching Filter Based on 3-D Sensor [5]:

As the equipment cost of the intuitive connection point of the huge screen framework is extremely high. Software for the design of a large portable screen has been developed in this paper. The main idea presented in this paper is that if an IR pen is projected on any plane, including the wall, it will produce infrared light. The Kinect profundity of the camera will then examine the intuitive wall, and the infrared sign will send it to the PC. First The PC will get the info information in view of exchanging fil- tering, and afterward control the intuitive programming for signal handling. The touch framework programming controls the intuitive procedure on the intelligent plane to empower human-PC communication Their framework utilized a 3D sensor (Kinect) to examine a projection surface. They did some ex- periments by squeezing the infrared pen switch, then, at that point, the infrared light coming from the IR pen was projected on the intelligent plane, and the Kinect sensor was called to catch the unfiltered picture. With the proposed exchanging channel strategy in this paper, the lengthy Kalman channel subsystem and the unscented.

3- Touch Sensing for a Projected Screen Using Slope Disparity Gating[6]:

This paper presents a touch-detecting framework for projected screens utilizing a solitary camera. Over- coming difficulties in picture handling and contact location, the framework utilizes Slant Divergence Gating to catch a particular locale over the projected screen. This area is intended to catch a piece of the finger when contacted and none when immaculate. Notably, the projector can both be used as an interface display and a programmable light source, allowing for selective imaging without changing how the user sees the projected image. The review incorporates a model framework, contact detecting calculation, assessments of touch exactness, and a showing of strength to different projected content. Figure 1: Comparative framework examination.

	Hardware	features	cost
Any-Wall Touch Control System With Switching Filter Based on 3-D Sensor	• Optoma PK301 DLP pico-Projector (resolution 854 × 480) • Logitech Pro 9000 camera(that provides a 960 × 720 image at 30FPS)	• Click/open any folder	$135
Any-Wall Touch Control System With Switching Filter Based on 3-D Sensor	• Infrared pen • 3D sensor (Kinect) • Depth camera • Small projector	• Click/open any folder • Write on the flat surface using the pen	$215
Touch Sensing for a Projected Screen Using Slope Disparity Gating	• Sony MP-CL1A laser scanning projector with 1280 × 720 resolution and 60Hz refresh rate • IDS UI-3250CP-C-HQ RGB camera, resolution 1600 × 1200.	• click/open any folder • Robust to lighting conditions	$1,034
Touch point detection (TPD)	Minimum requirements: • PC • Projector(optional) • Camera(no resolution limitation)	• Click/open any folder • Move cursor • Volume • Face detection	Free

Figure 1: Similar system comparison.

2.2 Business Applications

- **PapARt** [7] is a software development kit (SDK) that enables the creation of interactive projection mapping, papARt library is used by RealityTech Company for creating an interactive video game
- **Epson eb-695wi projector** is a device developed by Epson company that can be used as an input device to detect your hand movement on a whiteboard and transform it into actions the cost of this product is 80,000 EGP
- **Xperia Touch** A portable projector that transforms your wall or table into an interactive small-scale 23-inch touchscreen includes Sony's unique SXRD short-throw projection unit and strategically placed two-way stereo speakers for rich sound.

3 System Description

3.1 Problem Statement

The main problem that faces touch point projector devices is their high cost, making them unaffordable to many. Additionally, touch point projectors find it challenging to work effectively in different places, limiting their flexibility. Solving these challenges is crucial to ensure that these projectors can be purchased easily, and used in various environments.

3.2 System Overview

As shown in Figure (2), the system works as follows: the user connects his PC to the projector and then starts the camera. The projected content then is displayed on a flat surface for instance the wall. The camera gives the hand gesture commands and the users' facial features as input to the system. The system then detects the user's face using the YOLOv5 model, for the pre-processing we use contours detection to detect the borders of objects, also Gaussian blur is used to mute the image noise. Furthermore, Background Subtraction (BS) is used to separate the moving parts for example here we mean the user's hand movement from the background. Additionally, Histogram Equalization is used for image contrast optimization. Finally, for optimization techniques, threading enables concurrent execution of tasks within a single process, and multiprocessing enables parallel execution of tasks across multiple processes and is also beneficial for data-intensive operations.

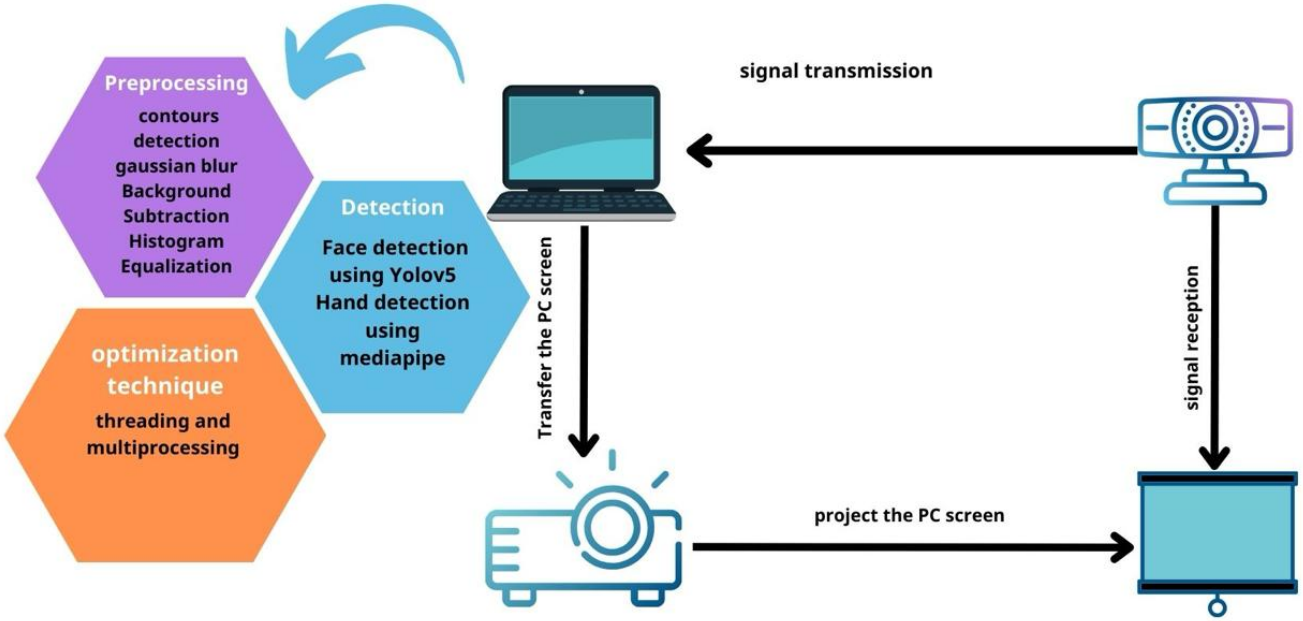


Figure 2: System Overview.

3.3 System Scope

The system will include:

- Gesture Recognition.
- Face Recognition.
- Hand Tracking.
- Single-Hand Interaction.
- customizing the actions for specific gestures.

And exclude:

- Multiple-Hand Interaction.
- Outside Frame Handling.

3.4 System Context

As shown in Figure (3), the TPD system has a PC that connects to the projector and a projector that displays the PC's provided content on a flat surface. A camera and an end user, the camera captures the user's commands and face features and sends it as input to the system, then the system analyzes it and reflects the given commands on the projected surface. Also, there is an account database that stores the account information of the user such as his credentials.

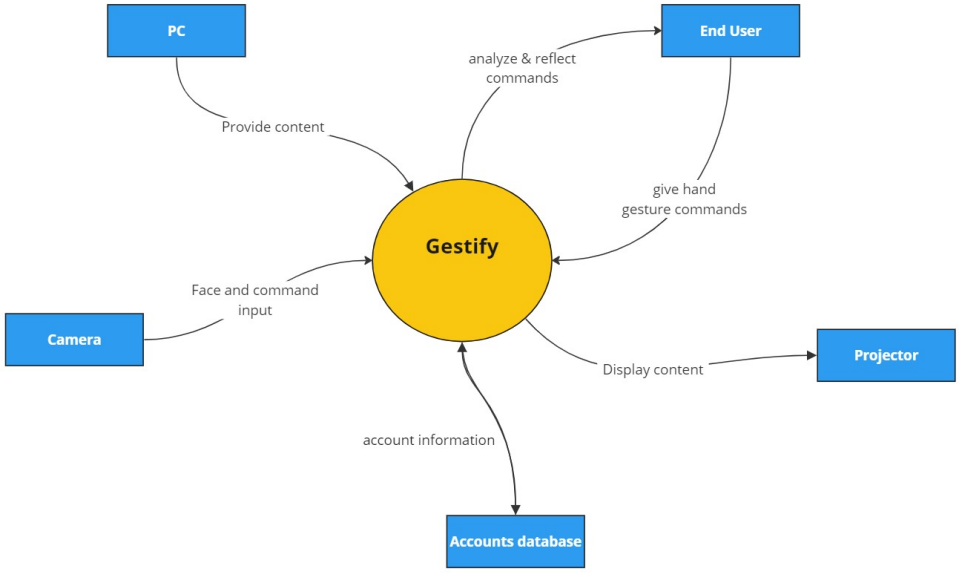


Figure 3: Context Diagram.

3.5 Objectives

Competitive price: Our optimal goal is the price. After making a competitive analysis we figured out that our system compared to others in the market costs way less than all enabling the users to run the system with their available hardware (PC-Camera-projector) with no hardware limitations at all, and more interactive features that are not available in any traditional touch projector.

Real-time Collaboration: Enable real-time collaboration by supporting interactions on the projected surface, fostering collaborative environments for business or educational purposes.

Low-Latency: The system must analyze gestures and reflect the corresponding action on the projected interface with very low latency.

3.6 User Characteristics

TouchPoint Detection users' characteristics will be as follows. Most users will be looking for an innovative solution to grab their audience's attention and engage them as much as possible. They will be either

- **Students** present a PowerPoint presentation to their professors at universities, and they need innovative yet affordable ways of engagement to deliver their content.
- **Professors** deliver lessons to their students, and they want to grab their students' attention.
- **Employees** presenting a new initiative or a business pitch.

To sum up, our users will be into more interactive, affordable solutions for their presentations, and that's exactly what the Touch-Point Detection Software was built for.

4 Functional Requirements

4.1 System Functions

The following figures represent the use case diagrams for the system. Figure(4) represents the Use Case diagram of the users interacting with the system, and Figure (5) represents the use case of the system itself.

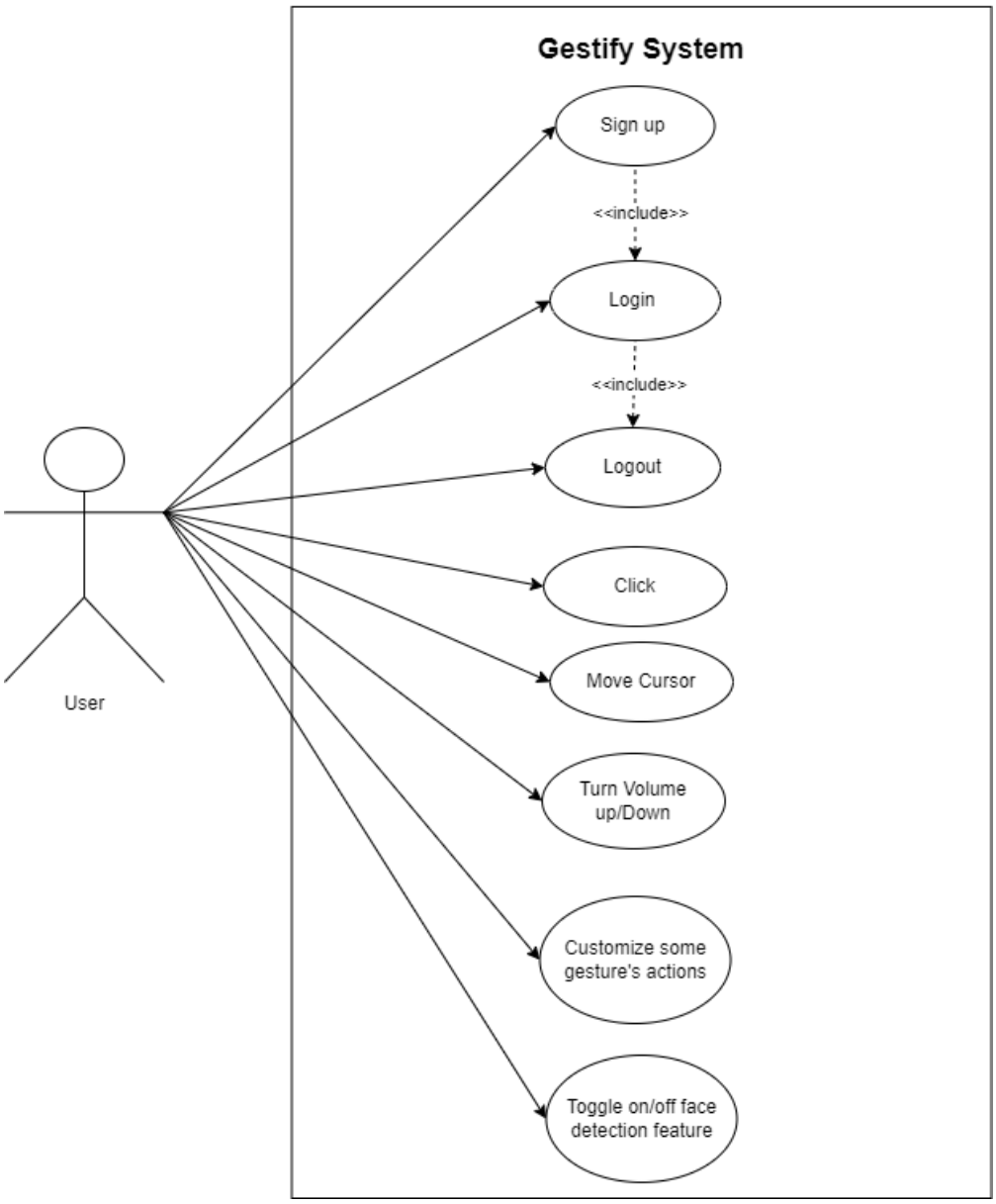


Figure 4: User use Case Diagram.

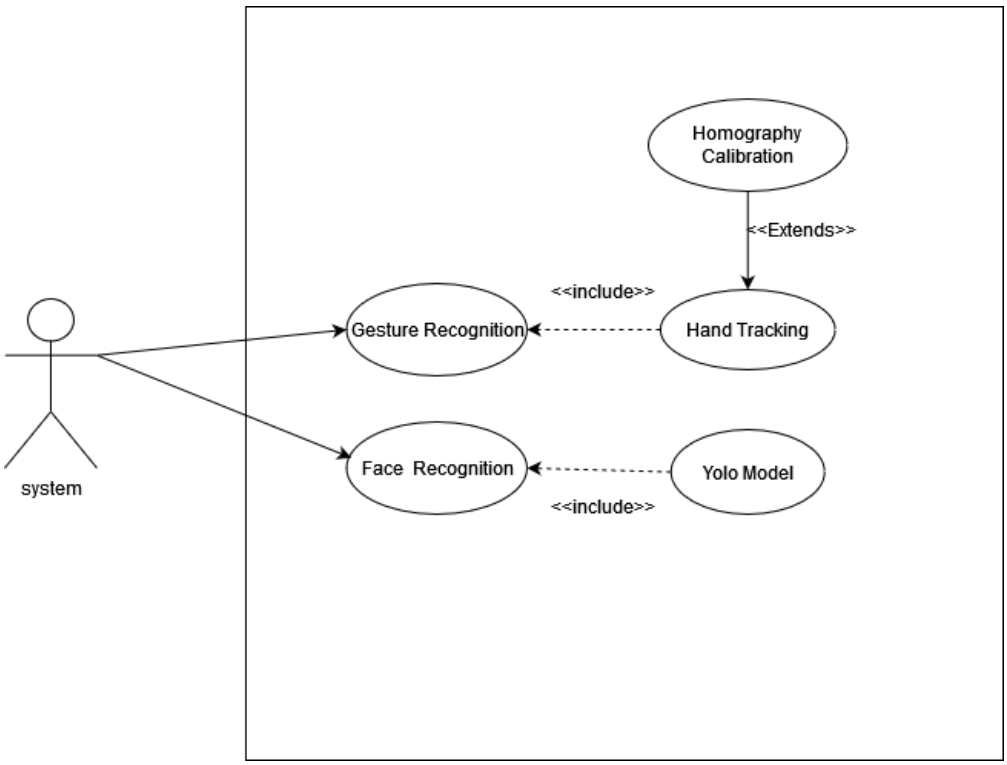


Figure 5: System Use Case Diagram.

Our key functions of Gestify Sofware are

- **ID:01 Sign up:** New users can create an account.
- **ID:02 Log in:** Current users can login with their credentials(email/password).
- **ID:03 Log out:** Users can log out from their accounts.
- **ID:04 Click:** where the user can open any file/folder - or click anywhere - by using his two adjacent fingers (the index and the middle). The initial step is to make a V-shape or peace sign with these two fingers and move them close to each other where there is no gap in between, then you will be able to give the click command and press wherever you like. This function is done using the media pipe library.
- **ID:05 Move cursor:** where the user's finger- for instance, the index finger- will be detected by the software using the camera, and can move the cursor anywhere inside the specified frame in the laptop.
- **ID:06 Turn up/down volume:** this function allow users to control their PC's volume by measuring the distance between their thumb and index fingers, all you have to do as an initial step is to make the 2 fingers(the thumb and index finger) touch each other, then expand the distance between them to turn up the volume. and to turn it down minimize the distance gradually.

- **ID:07 Customize some gesture's actions:** Customize the action of some gestures based on the user's preference through a drop-down menu.
- **ID:08 Toggle On /Off Face detection feature:** User can toggle on or off any time the face detection feature.
- **ID:09 Gesture Recognition :** The system will be able to recognize the user's hand gestures.
- **ID:10 Face Recognition:** The system will be able to recognize the user's face through its features, and this is done by using YOLOv5.
- **ID:11 Homography Calibration:** The system will be able to map the user's image pixel coordinates to real-world coordinates.
- **ID:12 Hand Tracking:** The system will be able to track user's hand.

4.2 Detailed Functional Specification

Name	Sign up
Code	ID 01
Priority	High
Critical	To create new credentials for new users so that they can log in next time easily
Description	This function is responsible for creating new accounts for users who don't have one
Input	Valid email address, password, confirm password
Output	Boolean: true if the password matches, and their email address was valid
Pre-Condition	User must install the software and run it, in addition, the user must not the same email and password saved in the database
Post-Condition	The user has to log in with his e-mail and password
Dependency	having a valid E-mail address and matching password
Risk	Signing in with an e-mail previously used in our system

Table 2: Sign up Function Description

Name	Face Detection
Code	ID10
Priority	Medium
Critical	To Make the system switch between user profiles automatically
Description	This function is responsible for switching between user profiles by his face features to get his custom gestures
Input	Click a button to activate
Output	Boolean: true for automatic account switching
Pre-Condition	Must be a registered user with his face stored
Post-Condition	Automatic profile switching
Dependency	Depends on already having a profile with the user's face uploaded
Risk	no profile is found or a bad user image

Table 3: Face Detection

Name	Gesture Recognition
Code	ID:09
Priority	High
Critical	To make the system work as intended
Description	It recognizes the movement of the hand and what action it performs depending on the gesture defined on the system
Input	Hand movement captioned through camera
Output	The intended hand gestures given by the user
Pre-Conditon	The hand should be in the area covered by the camera
Post-Condition	The gesture must be either already defined by default or custom assigned to specific action
Dependency	Depends on logging in to your profile either by logging in or by face detection
Risk	Getting out of the camera scope

Table 4: Gesture Recognition

Name	Customize some gesture's action
Code	ID:07
Priority	Medium
Critical	for making set of custom gestures for your profile
Description	Setting already defined gesture to specific action to assist you while using our system
Input	list of gestures to assign to actions
Output	full profile with your custom gestures
Pre-Condition	having an account
Post-Condition	having actions customized for some gestures
Dependency	depend on what set of gestures and actions does the system have and how you would like to assign them
Risk	not having an account

Table 5: Customize some gesture's action

5 Design Constraints

5.1 Standards Compliance

Gestify System is a cross platform software that runs on Windows,Mac, and Linux operating systems. Compatible with common projectors and cameras.

5.2 Hardware Limitations

The user must have a laptop with the software running on it, a camera -whether it's going to be the built-in laptop camera or an external camera like his mobile camera- to detect the user's face and hand gestures, in addition to a projector to present his content on a flat surface. Our software will run and give the typical features with or without the projector.

5.3 Other Constraints as appropriate

Simultaneous interactions from multiple users on the same projected surface is also a constraint Furthermore, Gesture Customization: users will be unable to customize all hand gestures, but will only be able to define an action for a specific gesture that is already defined in the system.

6 Non-functional Requirements

.

- **Performance:** Gestify system should analyze gestures/clicks and reflect the corresponding actions on the projected interface with very low latency, in addition to analyzing faces.
- **Reliability:** Minimize false clicks, false face detection, and false hand gesture errors during use.

- **Usability:** It should provide an easy-to-use simple interface for end users.
- **Compatibility with Hardware:** The software is compatible with common and most projectors and cameras.

7 Data Design

- For hand detection a Python library will be used called mediapipe which will be responsible for hand tracking and gesture recognition
- A dataset for face detection that contains 16103 picture label 393,703 faces with a high degree of variability in scale - (This dataset was exported via roboflow.com on October 24, 2023, at 7:30 AM GMT).[8]

  The following pre-processing was applied to each image:

  1- Auto-orientation of pixel data (with EXIF-orientation stripping)

  2- Resize to 640x640 (Stretch)

8 Preliminary Object-Oriented Domain Analysis

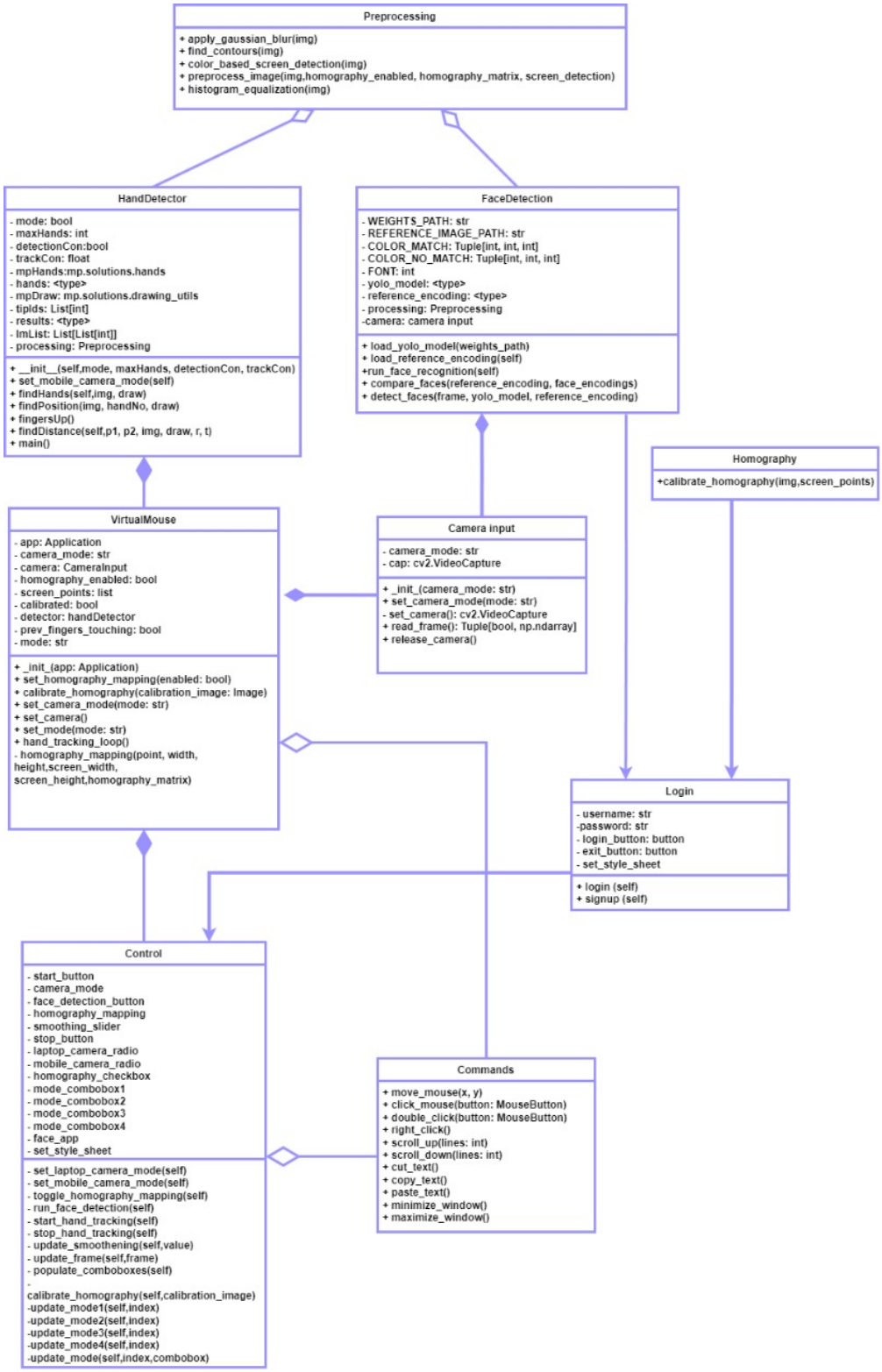


Figure 6: Class Diagram.

9 Operational Scenarios

Sign up scenario

- Initial assumption: A Student/Professor or an employee wants to start presenting a presentation or delivering a lesson, so they have their presentation on their computers ready to be displayed and TPD software is installed, but they don't have an account.
- Normal: The user then opens the software and clicks on the Signup button, and fills out the email address, password,confirm password fields correctly. Then he uses his credentials that have been just made to log in. On completion of the sign up, the system automatically saves his credentials to the database.
- What can go wrong: The user could write an invalid email address, in this case the system will notify him that he should use valid one The user could write a mismatch passwords, in this case the system sends him a message that passwords didn't match. Also, the user can sign up with already registered email address, in this case the system should notify him that this email is already used.
- System state on completion: user credentials is saved in the database and he can use them next time to log in.

Gesture recognition scenario

- Initial assumption: The user already logged in, run the software ,and presenting.
- Normal:User should be using the camera whether the build-in laptop camera or mobile camera. In addition using his hand gestures(default ones) to give commands(like: click,volume up/down,move cursor) within the camera scope.
- What can go wrong: The user could give commands outside the camera scope, the system will respond with a warning message on the screen that it can't detect his hand.
- System state on completion: The hand gesture will be recognized and the intended action will be accomplished.

Face recognition scenario

- Initial assumption: Two or more users already has an account and want to switch to their customized hand gestures. The already logged in user has already activated the face detection button,and there's sufficient light in the room.
- Normal: Users while presenting can switch to their predefined customized hand gestures by just standing in front of the camera and their face is recognized.
- What can go wrong:The already registered User's face won't be recognized by the system,in that case the system will send a warning message to him to show his face to the camera. for

the users who aren't registered in the system, the system will not do anything and keep the default/customized gestures as is.

- System state on completion: The system will use the customized gestures for the user that his face has been detected. And those gestures will be used to give their intended actions.

Customize some gesture's action scenario

- Initial assumption: The user already has an account and starts to customize his gestures by clicking the drop-down menu and choosing his preferred action accompanied by a specific hand gesture.
- Normal: The user after choosing his preferred action accompanied by a specific hand gesture,he should then save them. and start using those hand gestures specified by him.
- What can go wrong:The user may forget what hand gestures he chose and what action mapped to them -he will get confused- or he might click on the same action that should be taken for more than one gesture, the system here won't allow the user to choose an action he chose previously for another gesture.
- System state on completion: The system will use the customized gestures chose by the user for later interaction.

10 Project Plan

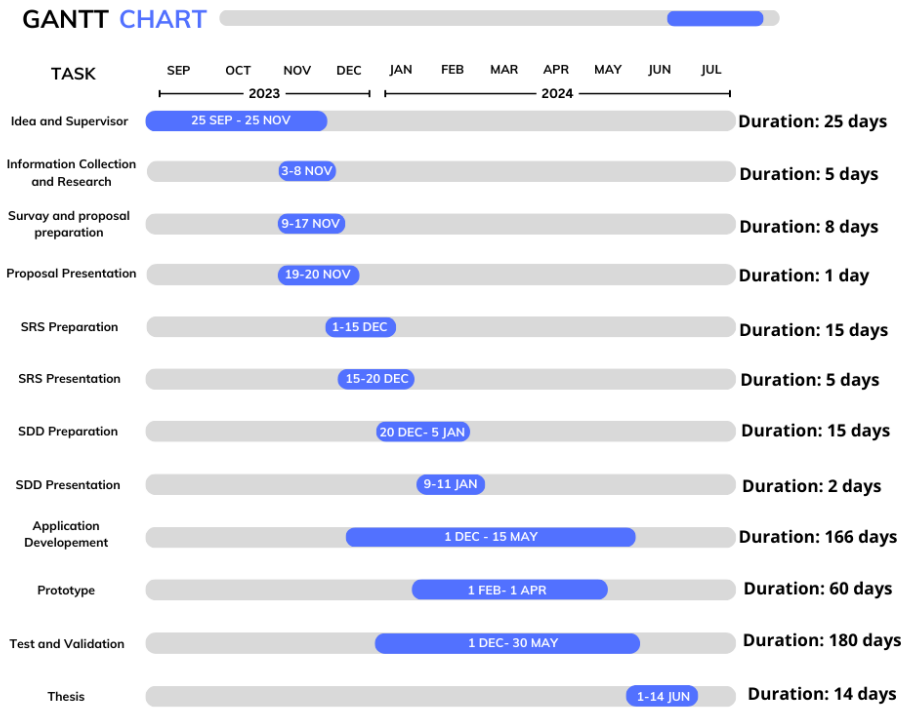


Figure 7: Gantt Chart Time Plan.

11 Appendices

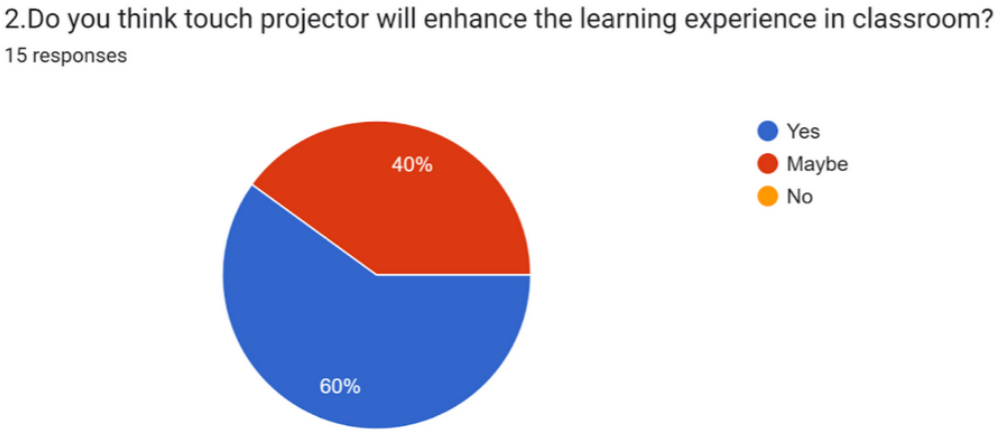
11.1 Definitions, Acronyms, Abbreviations

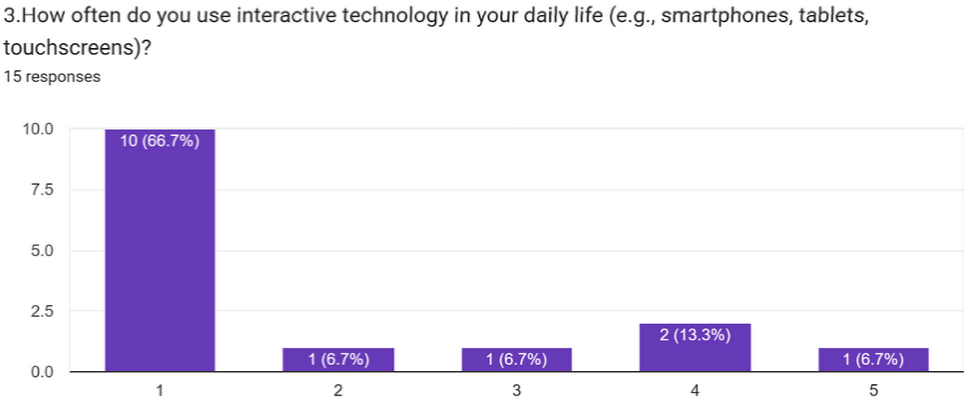
The following table will note the abbreviations' meanings used in the document:

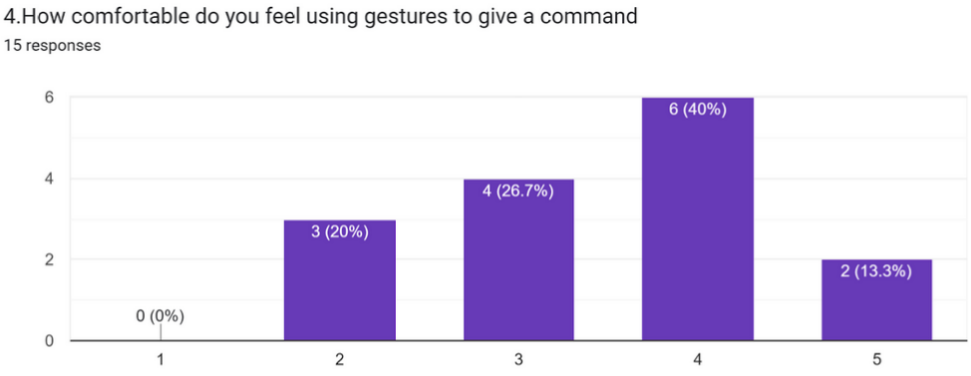
Term	Stands For
TPD	Touch-Point Detection
YOLO	You Only Look Once

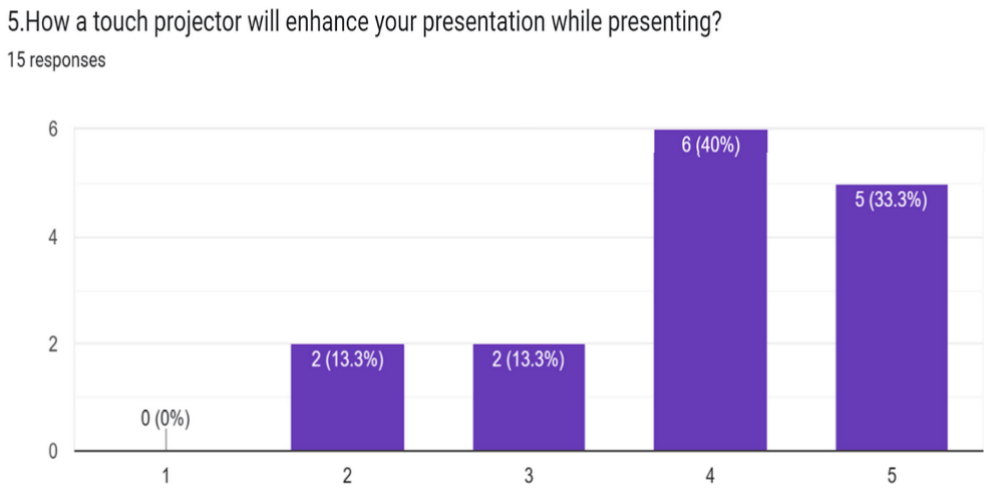
Table 6: Abbreviations Table

11.2 Supportive Documents









References

[1] Massimo Camplani, Luis Salgado, and Romolo Camplani. "Low-cost efficient Interactive Whiteboard". In: *2012 IEEE International Conference on Consumer Electronics (ICCE)*. 2012, pp. 686–687. DOI: 10.1109/ICCE.2012.6161809.

[2] Jingwen Dai and Ronald Chung. "Making any planar surface into a touch-sensitive display by a mere projector and camera". In: *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. 2012, pp. 35–42. DOI: 10.1109/CVPRW.2012.6239197.

[3] R. Kjeldsen, C. Pinhanez, G. Pingali, et al. "Interacting with steerable projected displays". In: *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition*. 2002, pp. 402–407. DOI: 10.1109/AFGR.2002.1004187.

[4] Christian von Hardenberg and François Bérard. "Bare-Hand Human-Computer Interaction". In: *Proceedings of the 2001 Workshop on Perceptive User Interfaces*. PUI '01. Orlando, Florida, USA: Association for Computing Machinery, 2001, pp. 1–8. ISBN: 9781450374736. DOI: 10.1145/971478.971513. URL: https://doi.org/10.1145/971478.971513.

[5] Lei Yu, Changdi Li, and Shumin Fei. "Any-Wall Touch Control System With Switching Filter Based on 3-D Sensor". In: *IEEE Sensors Journal* 18.11 (2018), pp. 4697–4703. DOI: 10.1109/JSEN.2018.2827386.

[6] Mayuka Tsuji, Hiroyuki Kubo, Suren Jayasuriya, et al. "Touch Sensing for a Projected Screen Using Slope Disparity Gating". In: *IEEE Access* 9 (2021), pp. 106005–106013. DOI: 10.1109/ACCESS.2021.3099901.

[7] Jeremy Laviole and Martin Hachet. "PapARt: Interactive 3D graphics and multi-touch augmented paper for artistic creation". In: *2012 IEEE Symposium on 3D User Interfaces (3DUI)*. 2012, pp. 3–6. DOI: 10.1109/3DUI.2012.6184167.

[8] Large Benchmark Datasets. *WIDER FACE Dataset*. https://universe.roboflow.com/large-benchmark-datasets/wider-face-ndtcz. Open Source Dataset. visited on 2024-01-08. Oct. 2023. URL: https://universe.roboflow.com/large-benchmark-datasets/wider-face-ndtcz.