Sheffield Hallam University | Faculty of Science, Technology and Arts

MSc Dissertation Report

Implementing Facial Recognition Technology Using a Microprocessor to Register Student Attendance

A dissertation submitted in partial fulfilment of the requirements of Sheffield Hallam University for the degree of Master of Science in Information Technology Management

Student Name	James Early
Student ID	28040066
Supervisor	Dr Elizabeth Uruchurtu
Date of Submission	16th September 2019

This dissertation does NOT contain confidential material and thus can be made available to staff and students via the library.

Abstract

This report explores the attempt to discover a deeper understanding into the consistency behind facial recognition systems. The focus is to provide enlightenment into how they operate within the boundaries of today's devices. An attempt to transfer the accumulated knowledge to solve an identified real-world issue will also be studied. Specifically, an in-depth investigation will be carried out to examine how educational institutions can benefit from employing such systems.

A review of expert literature will be performed to expose any scenario in which this type of technology has been applied previously. This will then allow for a comparison to be made by analysing the different architectures to recognise the optimum solution to be applied within an educational environment. This research could potentially be used to better the existing solutions as well as expanding the horizons of using facial recognition in different settings.

The depiction into how the facial recognition technology performs will be illustrated through the medium of a Raspberry Pi Microprocessor. A prototype will be created containing the ability to execute various functions through the use of a Python interface. The focus will be on using participants to simulate a classroom environment that allows for effective testing of the chosen facial recognition algorithm and software libraries.

Due the nature of research being conducted, ethical consideration presents a substantial footprint throughout the project therefore methods in which to properly manage it will be expressed. Data security is paramount as sensitive information will be stored within the prototype itself. Suitable methods will be identified to establish effective ways to reduce the risk of data leakages thus preventing law infringements or breaching data protection acts.

Finally, a conclusion will be drawn to analyse the overall performance of the prototype system. A discussion will take place to determine if the research question has been answered in addition to if the aims and objectives have been met. Limitations will also be articulated along with any potential future development plans for the project.

Acknowledgements

A tremendous thank you and appreciation goes to my Supervisor Dr Elizabeth Uruchurtu who has provided the support during not only her time on this project, but also from the initial stages of the course. With her consistent supervision and knowledge, perseverance has been maintained to see the project through to its conclusion. Thank you again!

One more sincere thank you goes to my close peers, Suganthan and Mike. We were continually there to motivate each other throughout the entirety of our studies.

Dedication

This project and report will be devoted to my mother and stepfather. Through their unconditional financial and psychological assistance, they have given me the confidence to achieve my personal and academic goals.

I'd also like to add my recently passed Grandmother. Though just missing out on her 100th birthday by a matter of weeks, she was always there to make me laugh and provided the most important advice one could give - *"Life is short. Live every day like it's your last"*

Keywords

Facial Recognition, OpenCV, Deep Learning Metrics, Raspberry Pi 3 Model B, Microprocessor, Python, Student Attendance System

Contents

Abstract i
Acknowledgements ii
Dedication ii
Keywords ii
Chapter 1 - Introduction 1
Chapter 1.1 Project Concept and Research Problem 1
Chapter 1.2 Research Question and Objectives 2
1.2.1 Research Question 2
1.2.2 Aims and Objectives 2
1.2.3 Deliverable 3
Chapter 2 – Literature Review 3
Chapter 2.1 Review 3
Chapter 3 – Methodology 12
Chapter 3.1 Data Collection and Research Design 12
Chapter 3.2 Performance Measurements Indications 13
Chapter 3.3 Ethical Considerations and Security 15
Chapter 3.4 Resources and Participation 17
Chapter 4 – Implementation 19
Chapter 4.1 Front and Back-End Design 19
Chapter 4.2 Recognition Algorithms Comparison 25
Chapter 4.3 Testing Environments 28
Chapter 5 – Evaluation of Results 34
Chapter 5.1 Expectations Vs Results 34
Chapter 5.2 Performance Analysis 40
Chapter 6 – Discussion 43
Chapter 6.1 Research Question Re-evaluated 43
Chapter 6.2 Limitations 45
Chapter 7 Conclusions and Further Work 46
References 48
Appendices 50
Appendix A – Research Project Plan 50
Appendix B – Ethics Checklist Form 57
Appendix C – SHUREC8 63
Appendix D – Data Management Plan 69

Appendix E – Ethics Committee Approval....71
Appendix F – Information Sheet....72
Appendix G – Consent Forms....75
Appendix H – Results of Exercises....79
Exercise 1 – Still Image Comparison....79
Exercise 2 – Live Video Feed Detection Results....79
Appendix I – Python GUI Code....80
Appendix J – Facial Recognition Script (Still Image Comparison Exercise)....83
Appendix K - Facial Recognition Script (Live Video Feed Exercise)....86
Appendix L – MySQL Database Creation Script (Exported)....91
Appendix M – GANTT Chart....93

Chapter 1 - Introduction

Chapter 1.1 Project Concept and Research Problem

The goal for many schools and universities across the country is being able to reduce the absenteeism of its students. (Epstein, 2010) The issue is mainly related to the method in which their attendance is being captured. The most traditional approach is using the basic technique of pen and paper however in some more technologically advanced institutes, ID card scanning is being utilised. Marking a student's attendance is incredibly important as it provides staff an overview into their overall presence.

Current methods such as paper-based recording is still used as the primary technique in most education institutions yet pose a substantial amount of disadvantages over other available means. The most evident would be the issue of maintaining an accurate representation of who's present especially in a large classroom. Many lecturers and teachers circulate a single sheet of paper that requires the students to sign next to their name which in itself is a lengthy process and has the potential to cause a disturbance to an already started lesson. The probability of each student in the class marking their attendance successfully is low particularly in a room type such as a lecture theatre or auditorium. Moreover, the paper-based register that is passed around is exposed to damage or even has the potential of getting misplaced. (Jha, 2015)

It seems that the issues identified above share a common theme that due to human interaction, traditional student attendance systems still pose a wide range of problems. If a system was implemented that removes the need for participation, then a better robust solution could be used. The ideal type of technology to investigate would be the use of biometrics as it utilises the body for identification. Biometrics is being integrated into devices such as mobile phones but has yet to be seen within an education institution. Compared to the visual attendance systems of ID cards and student signatures, biometric detection has the potential to be a much faster and efficient method to record attendance. (Alterman, 2003) It also requires minimal effort from a user's perspective thus solving the disturbance issue during classes. (Kar, 2012)

By combining existing technologies such as a microprocessor and facial recognition would allow for the construction of a unique system. Implementing a new system that requires nominal management and is automated will eradicate any possibility of producing inaccurate attendance data. The current process being used by Sheffield

Hallam University includes students manually recording their presence via a signature which poses many potential issues such as missing out a student or even forgery. This only emphasises the fact that a new system arrangement is needed.

Though providing an automated device is pragmatic, the important concept of ethics needs to be examined. When enforcing any form of biometric identification, raises a significant amount of concerns regarding privacy especially when capturing and storing information about those exposed. (Alterman, 2003) The concern mostly revolves around the security and usage of the data as recently, there have been cases where large organisations have lacked ethical professionalism and breached data protection acts. The ethical considerations of biometrics cannot rely on the assumption that the data is secure therefore a discussion will be made regarding the possible ethical risks involved in the project.

The underlying aim is to create a process that is capable of capturing student's attendance in class. It needs to be reliable enough to supersede the current architecture and possess the ability to produce a full-bodied set of results that can be used to monitor student's presence. It's important to realise that this project is essentially building a pragmatic method into solving a research problem through the medium of technology rather than investigating the wider issue of generic student registration systems. It's imperative to understand that the outcome of this project will solely consist of a prototype device therefore it cannot be emphasized enough that it will not be implemented in a live classroom environment.

Chapter 1.2 Research Question and Objectives

1.2.1 Research Question

How to quantify a student's attendance to produce accurate registration results?

1.2.2 Aims and Objectives

- Build a Raspberry Pi prototype with encompassing LCD screen and USB camera attachment
  S – Compact system that allows for easy demonstration.
  M – LCD screen isn't essential but will provide a tidier deliverable.
  A – The resources are cost-effective including the software development IDE which is free.
  R – Essential to the projects overall success as it provides the computing power.
  T – Initial objective to accomplish as this will set the foundation of the project.

- Implement OpenCV and Deep Learning to enable facial recognition functionality
  S – Install the required pre-requisites to permit the use of facial recognition.

M – Both are required to output the results needed.
A – Precise modules and packages need to be installed in the correct order to avoid corruption.
R – Essential to the projects overall success as it provides the facial detection.
T – Required to be completed before GUI creation.

- Develop a Python GUI that allows for ease of use when operating system
  S – Provides a simple GUI to execute CMD/terminal functions.
  M – Enough to fully test the system.
  A – Python knowledge needs to be applied to created adequate GUI.
  R – Non-essential however extremely pragmatic.
  T – Cosmetic objective therefore can be completed towards the end of project lifecycle.

- Output a set of feasible results that could be later used for analytical purposes
  S – Export results gathered using MySQL tools.
  M – Simple relational database to display basic results of attendance.
  A – Prototype local database will be implemented to avoid hosting costs.
  R – Added practicality to the project.
  T – Minor time to build therefore can be built towards the end.

1.2.3 Deliverable

The deliverable will be a fully functional Raspberry-Pi facial recognition system that has the ability to accurately simulate and display the results of a student's university class attendance.

Chapter 2 – Literature Review

Chapter 2.1 Review

An imperative element for the success of the education system is for the regular attendance of its students. This is inclusive for most higher education institutions however it can still be applied to lower level schools. One of the main issues today is getting students to attend their classes. (Epstein, 2010) This is becoming increasingly more frequent with university's as being present is no longer compulsory. This is a complete contrast for younger students who attend nursery and high-school as it's seen to be a requirement more than an option. That being said, many university's still have a benchmark of attendance expected from all students in the form of a percentage. If this isn't reached, then disciplinary action could occur.

This alone is still not enough to force students to attend therefore a certain degree of stimulation for a student needs to be applied. Traditionally, students should naturally understand that maintaining a constant presence in the majority of their classes will

constitute to their overall grade. (Krishnan, 2015) In some scenarios, schools have started to use attendance points to promote the appearance of students which can then be used to reward or penalise them depending on how many they accumulate over a period of time. This method seems useful but still possess the risk of how the attendance is recorded.

As discussed before, most modern systems manually record attendance either by using a paper-based register or scanning ID cards. This has proven to be a pragmatic approach as we still utilise this system today however it does contain its own set of issues. The most apparent drawback is that it adds an additional pressure onto the teacher to correctly mark the attendance of the present students which consequently also has the ability to disrupt the class. (Lukas, 2016) Other issues could include the possibility of missing out students entirely especially in a large classroom environment. (Krishnan, 2015)

Even though there are multiple problems relating to the current arrangements being used, this project focuses more into how we can record the data more accurately. By creating a robust method into capturing this type of data will eliminate any previous or subsequent issues identified with existing techniques. There are a magnitude of ways to capture the attendance of pupils whether it's using traditional methods or integrating technology. The most underused technique would include the incorporation of biometrics especially within an education institution. Biometrics such as facial recognition possesses the merits of both accuracy and low intrusion which would make it a perfect solution to replace current data collection methods. (Kar, 2012)

However, implementing this type of technology within any type of setting will raise a substantial amount of questions involving ethics. Having the ability to scan a student's face without consent or their awareness presents a range of issues regarding their own personal privacy. It may also lead to students questioning the usage of their personal data. Especially in recent times, we have seen multiple large organisations such as Facebook undergo scrutiny and investigation into how its user's data is managed which could potentially repulse today's students from accepting the proposed system.

Both historically and in modern times, ethics has always been a major consideration when identifying methods in which to handle its user's information. This occurs within

the majority of organisations whether it's the client's data or its own staff members. This is even more so the case when working within the IT industry as the occupations predominantly involve working with someone else's data. The concept of facial recognition technology is controversial as it's centred around the idea of capturing someone's identity. Though the efficiency of implementing it in an educational environment seems ideal, the ethical consideration needs to be taken seriously and managed appropriately.

Due to the extensive capabilities facial recognition has to offer, we are now seeing it being implemented in public spaces across the world. (Brey, 2014) The main focus is to prevent theft and identify individuals within certain high crime areas. It's also being integrated into SMART CCTV systems in areas such as shopping malls and high streets. (Brey, 2014). Applying this type of technology to prevent potential threats and maintain public safety seems like a logical approach however this also comes with a number of ethical issues. The SMART CCTV cameras are required to collect information about members of the public before the facial detection can function. Inevitably, many people will be dismayed once they discover that their personal information is being collected without their consent. (Nissenbaum, 2010)

Knowing that your data is being collected, stored and analysed will not only raise personal privacy concerns but also question how the data is used. Function creep is becoming a frequent affair for facial recognition technology as there is no clear definition or legislation for its use. (Brey, 2014) Alterman believes that function creep may even extend into third party company's purchasing data to be used for marketing purposes. (Alterman, 2003) This raises a valid point as incidents similar to this aren't uncommon as we've already identified the provocative case of Facebook and Cambridge Analytica misusing and selling users data. (Cadwalladr, 2018) Brey proposes that if detailed legal standards were defined regarding proper use of facial recognition technology in pubic places, the issue of function creep will be significantly reduced as third party company won't be able to purchase personal information. (Brey, 2014)

After the terrorist attacks on September 11th 2001, biometric identification was immediately the focus on ways to prevent further incidents. The approach taken was to verify and identify individuals in key areas such as airports. (Bowyer, 2004) The

study on implementing facial detection became a success allowing it to be integrated into ‘facecams’ equipped by police officers. This was again to prevent crime through the use of biometric identification. At the time, applying this technology increased the productivity and decreased the crime rate in certain areas but it still provided inaccurate readings causing innocent people to be arrested. (Brey, 2014)

This was especially the case when implementing the technology within CCTV cameras in shopping malls as due to the lack of ingenuity present in facial recognition algorithms, many members of the public were being falsely accused as a potential criminal leading to harassment by the police. It was obvious that society relied on the technology too much as shopping malls only had a positive detection rate of just 69%. (Alterman, 2003) Since then, the efficiency and versatility of facial recognition has significantly improved however the issues presented by invasion of privacy and ethical professionalism still cause outrage amongst many thus creating the argue of, are the public willing to accept the trade-off between their own personal privacy for the safety of others? Brey believes this as he expresses that the public tend to accept the minor inconveniences so that criminals can be apprehended. (Brey, 2014)

For student attendance, applying mechanics such as biometrics would prove to be much more effective however it does come with limitations. Implementing a system that utilises facial detection will warrant suitable hardware that is powerful enough to capture faces real time. In conjunction is also choosing an algorithm that is simple enough to work alongside the CPU without overloading it. Proposed facial recognition systems usually include multiple algorithms and adequate computing hardware capable of producing accurate readings from the students. “We have taken different techniques like colour based detection and Principle Component Analysis (PCA) for face detection and for feature extraction, PCA and Linear Discriminate Analysis (LDA).” (Jha, 2015)

We can see from Jha’s implemented design that he uses the popular PCA and LDA algorithm that extracts the colours and features of the students face for increased facial detection results. This solution is the perfect answer for the research question however having the resources to execute and maintain this architecture will be challenging. Similarly, another proposed system uses the AdaBoost algorithm and SOC hardware

framework to carry out their facial recognition (Fuzail, 2014) but again, implementing this setup using the resources available will not solve the complication identified.

The most substantial challenge of the project is to build a prototype that is able to solve the issue whilst utilising the resources accessible. We have previously seen the requirements into how to get accurate results using high specification components however the difference is achieving the same results using lower level assets. In this instance, using a Raspberry-Pi microprocessor along with basic facial recognition algorithims such as the Haar Cascade and HOG detection provided by OpenCV and dlib libraries. They may not produce the highest results which would be achieved by using either Jha's or Fuzails implementations however it should still solve the ongoing issue that current student attendance systems present.

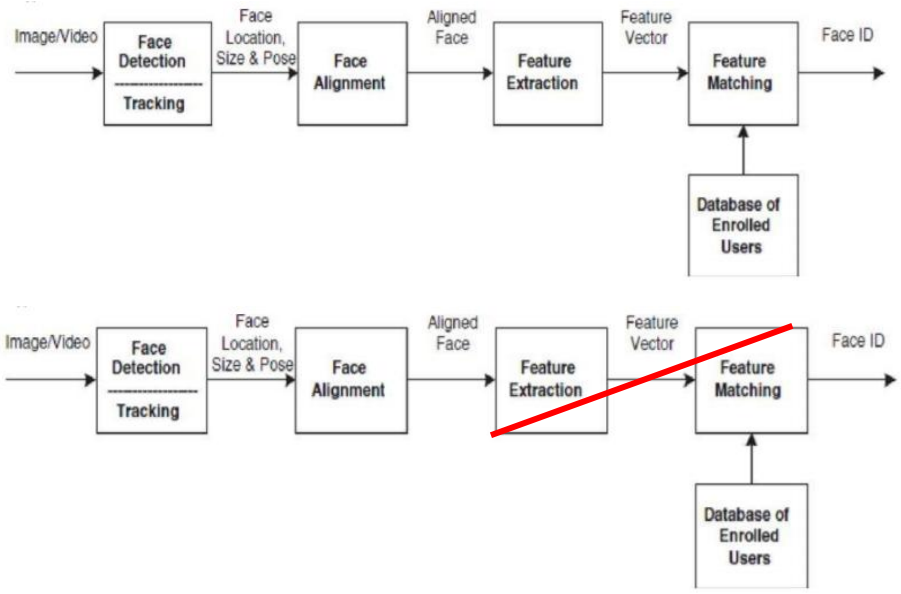


Figure 1 *- Comparison of Jha's algorithm to the projects prototype (Jha, 2015)*

Above is an illustration into the comparison between the algorithm that will be implemented to Jha's. Due to the processing capabilities between a microprocessor and Jha's architecture, feature detection will need to be disabled. Jha's system also incorporates the need to create a report for the student which is feasible however doesn't add any real value for this project. We can see his approach below.

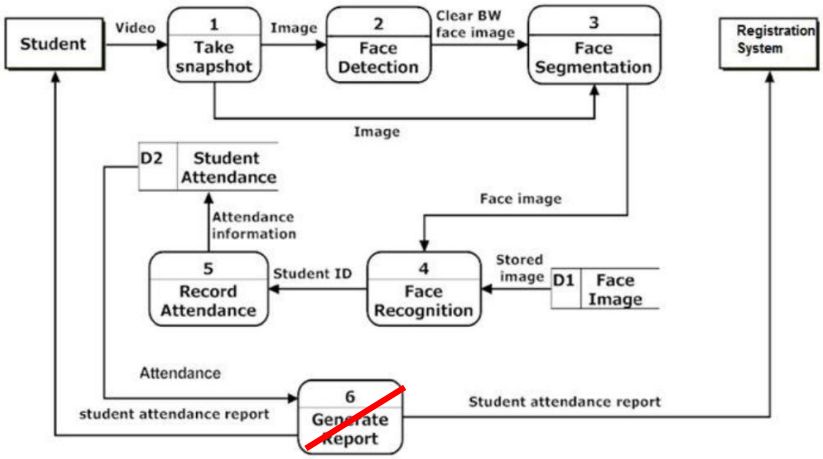


Figure 2 *- Activity Steps of Jha's Solution (Jha, 2015)*

To give an idea into the architecture comparison, below you can see how the resources available differ from Fuzail's. Whereas the proposed solution contains all components being hosted on one device, Fuzail's system contains multiple machines that each share the processing load. For this project to be a success, there needs to be the correct balance between choosing the correct algorithms that work effectively with the chosen hardware all while producing accurate results.

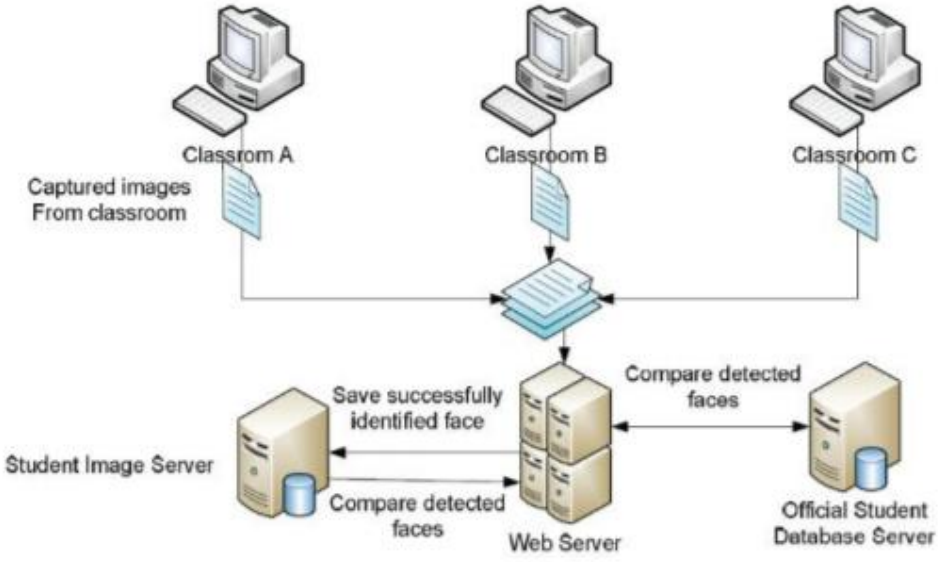


Figure 3 - *Fuzail's hardware architecture (Fuzail, 2014)*

The next solution provided by Sajid incorporates the usage of 2 databases. One contains the images of the students and the other the attendance records. (Sajid, 2014)

The databases are both created in an SQL environment allowing for manageable data storage. Although Sajids student images are saved within a SQL database, this projects are stored on the devices hard drive leaving it vulnerable to both cyber-attacks or corruption due to potential hardware failure.

Sajid also explains that his system has the ability to randomly take the attendance three times during a lecture or seminar. (Sajid, 2014) The idea of this feature is to ensure that the students who are enrolled in the class are still present after the lecture has commenced. This avoids the issue of students who may arrive to then immediately leave for the sole purpose of being marked present thus evading absenteeism punishment. Employing this capability creates a truly unique system and provides improved attendance results validity. When being compared to the proposed solution, there is a substantial hardware limitation as running the facial detection algorithm more than once during a class runs the risk of overloading the processor causing it to overheat and possibly hang.

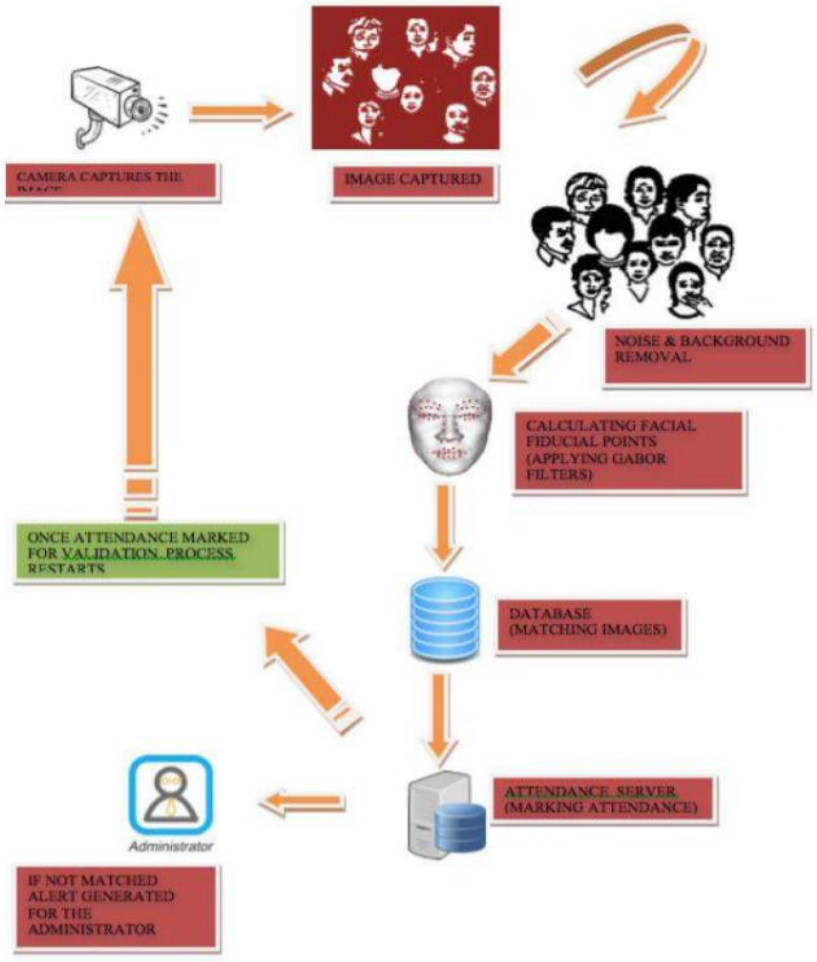


Figure 4 - *Activity steps of Sajid's solution (Sajid, 2014)*

The final solution discovered is one that exploits the credentials of the student's university logins. Once the students enter the classroom, they are required to input their username and password in order turn on the camera. (Rhesa, 2014) Once all have entered and signed in, the facial recognition function will be enabled. In essence, each sign in electronically creates the classroom inside the database allocating each student to that particular class. This strategy is effective as when new students are enrolled onto the course, the database can be adapted on demand and will only validate those who attend on that day. However, this approach will inevitably prolong the attendance process rendering this method impractical.

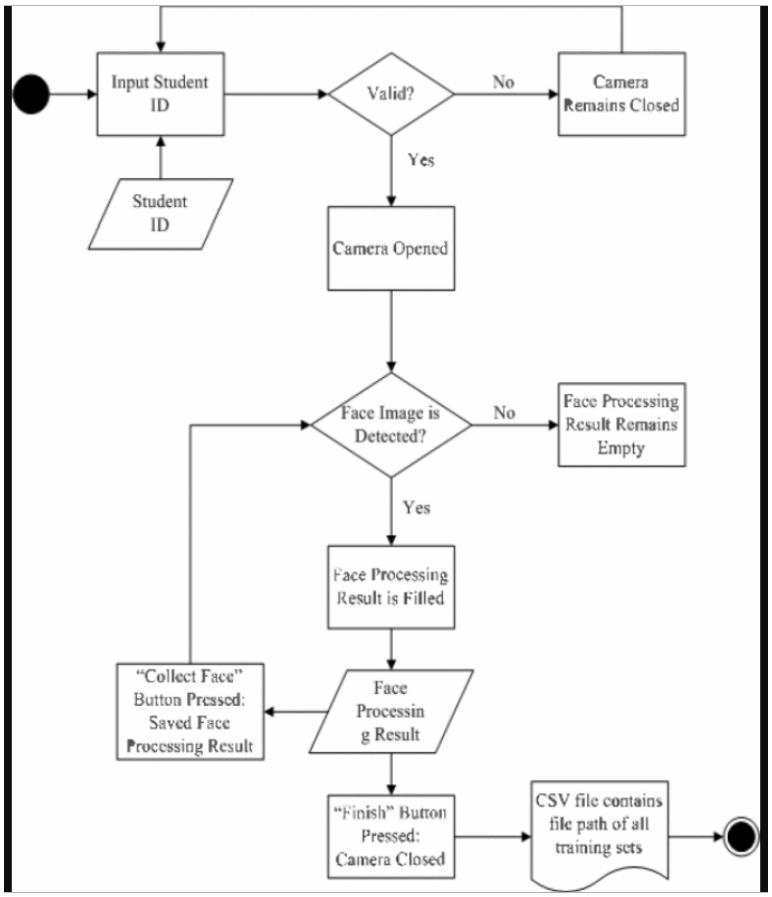


Figure 5 - *Activity steps of Rhesa's solution (Rhesa, 2014)*

Analysing each existing system applied in an educational environment provides insight into different avenues when trying to capture student's attendance. It's important for this project to examine each approaches advantages and disadvantages to build one that is capable of solving the research question. It's evident that the main issue is regarding the available resources of the prototype. Many of the existing systems have more than one machine to share the processing load therefore occupy the ability to include more complex features and efficient algorithms thus producing improved results.

The most suitable adaptation would be to use Sajid idea of implementing 2 databases which separates the student's images from the attendance records. Fuzial solution also presents the notion of dividing each set of students into separate classrooms allowing for easy data management. Jha's provides a more statistical approach by providing the idea of creating reports to illustrate the student's overall attendance data. Taking all solutions into consideration, the hardware design and activity diagram of the proposed system is as follows:

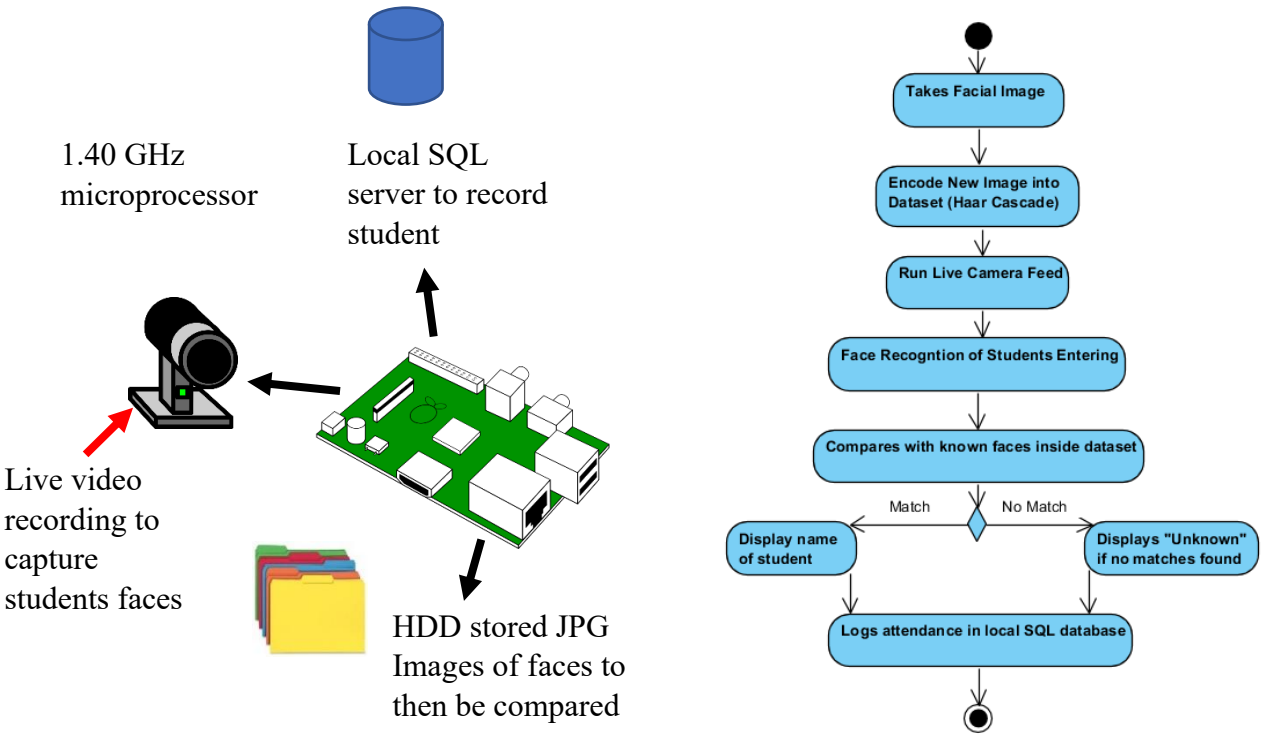


***Figure 6** – The projects proposed solution and activity steps*

Chapter 3 – Methodology

Chapter 3.1 Data Collection and Research Design

Choosing the correct type of data in conjunction with the most suitable method in which to capture it is paramount to the projects overall progression. It's vital to appreciate that the consistency of data being managed holds a substantial threat to the participants confidentiality therefore recognising a safe and secure process in which to extract it will aid into preventing ethical issues. The data also needs to be adequate enough to enable extensive testing of the prototype. The technique of accumulating the data will involve taking a clear image of the participants faces. This could be either from the USB camera attached to the Raspberry Pi or from their social media accounts.

The participants will be given a choice into how they would like their image captured. With consent, images from their social media profile will be used. This will provide accurate facial recognition readings however using this approach will compromise the authenticity of results. Those who take part in the study will be encouraged to use the USB camera as the data captured will articulate the efficiency of the prototypes performance when executing different functions.

The project itself will be built on the foundations of quantitative data as the majority being managed will include images of the participants faces. Using this type of data allows for the optimisation of results over other types such as qualitive. The images collected will act as the fundamental information that allows for the project to operate and produce the outcomes expected. The outputs of the project will also be quantitative as observations will be made to assess how the prototype performs.

The characteristics of the data will be in the form of standard JPEG images as these will be suitable for extensive facial recognition testing. The data gathered can then be used to determine the effectiveness of the system and provide an insight into the accuracy of the algorithm employed. The reliability of the measurements are crucial as these could be used as future work into improving the chosen algorithm or find alternative ways into generating improved results. As an outcome of using the quantitative design, utilising the mono-method will be a suitable technique for the collection and analysis of the data.

It's important to adopt the most appropriate philosophy for the project as it provides a sense of realism. This is usually shaped from previous experience and knowledge

which will constitute how the research question is interpreted. The most adequate philosophy that can be delivered is the one of a pragmatist. This is because the data being collected will effect the actions that lead into answering the research question. This will only be the case if the gathered data is credible and precise therefore by extracting clear images of the participants faces will mould the rest of the projects infrastructure thus producing the results expected and potentially revealing an answer to the research question.

Also, there is no singular viewpoint that can be large enough to portray the bigger picture. As identified before there are alternative student attendance systems that exist hence in another reality, approaching the project differently could potentially provide more ideas into improving the proposed system. This state of mind benefits the project over other philosophies such as realism. This is mainly because this mentality largely revolves around large datasets and includes qualitive data which the project doesn't contain.

As the project isn't part of an existing case study, it's imperative that the research strategy implemented caters for what it's trying to achieve. We already have acknowledged that it isn't building a new solution but rather adapting current architectures to solve a specific problem. Consequently, adopting the feasibility study is ideal as it's oriented around the idea of taking an existing system and discovering if it's possible to make it work with the resources available. This strategy allows for the creation of a prototype which should be built around efficiency using the quantitative data gathered.

Chapter 3.2 Performance Measurements Indications

Completing the project requires extensive testing of the prototype's functionality. The main focus will be on assessing the performance of how efficiently the facial recognition algorithm runs using a microprocessor. The accuracy will also be examined as a system with less CPU processing capabilities may produce inaccurate detection results. It's important to understand that microprocessors such as the Raspberry Pi have substantially fewer resources to distribute therefore it will be interesting to observe the effects this has on the outcomes of the project.

The preliminary stage is to store the participants images within the file directory and encode them into the dataset. These are then used as the source images for comparative

purposes when being analysed against any others during the study. The dataset acts as a database that stores all facial recognition information for each participant. After this stage is complete, the prototype will be equipped with the required data for the testing phase.

An initial indication into the extent of the prototypes capabilities would include observing the accuracy of the facial recognition algorithm on a still image. The optimum approach would be to observe how the system performs under this condition and test the detection rate. This data can then be analysed to see in exactly what scenarios the identification works greatest. To establish the best outcome possible, an array of different measurements will ensue to gauge the efficiency of each function the prototype has to offer.

The main focus of the protype will be to observe the accuracy of the facial detection on both a still image and during a live camera feed. This will be achieved by recording if the algorithm chosen will correctly recognise each participant successfully and maintain detection consistency. This is especially important when running the live video feed as it's vital that the system upholds and displays the correct identity of the participants for the duration of the exercise. However, there could be issues when testing the system on individuals who share similar facial characteristics as incorrect identifications could be made. If there are participants who do appear to be alike, observations on this aspect will also be made and recorded.

To further extend our knowledge into the depths of the prototypes abilities, is to assess other technical facets. Another essential area to examine would be the maximum distance in which the system can successfully detect individuals. It's crucial that the camera can scan faces from a reasonable distance as in a real classroom environment, it's unknown where it will be positioned. However, there are multiple variables that may affect the systems aptitude to scan from a distance. For instance, the hardware quality needs to adequate in order to fully utilise what the facial recognition software offers. Specifically, the USB camera needs to support high definition output in order render any faces detected. Therefore, a 1080p or 720p HD camera will be appropriate for long range identification.

Running the system in an ideal setting won't always be the case. In many situations, facial recognition will need to perform when being exposed to environments that will

challenge its potential. Atmospheric changes will inevitably occur therefore testing the efficiency in these instances will provide clarity into how well the prototype operates overall. The most likely to happen is the sudden change in lighting. Typically, classrooms have a variety of lighting conditions with some being darker or brighter than others. Tests will be carried out to observe if the accuracy drops when trying to detect faces in a reduced lighting setting.

Once the optimum scenario for capturing data has been identified, the simulation of a live classroom environment can commence. The results produced from this exercise are imperative towards revealing a potential answer to the research question as it will exhibit the full capabilities of the protype system. The experiment will consist of seeing the participants to walk past the camera and observe the consistency and accuracy of successfully identifying each individual. A possible issue could include severe latency during the live feed as using a microprocessor may not provide the camera enough resources in order to maintain a high frame rate.

Chapter 3.3 Ethical Considerations and Security

Carrying out this type of project eluded for the possibility into solving an identified issue however it also possesses its own risks that prevent it from completion. We have discussed previously the type of data being captured but haven't deliberated the potential impacts it could cause on both the participants and society. The presence of ethical consideration holds a substantial footprint throughout the duration of this project and needs to be examined before any sensitive data is extracted.

Information recorded from the participants will include compromising images of their faces. This alone has the potential to expose their identity to anyone who comes across this data. Taking the correct steps to prevent any being mistreated must be rigorous in order to maintain ethical professionalism during the project. We have seen previously that ethical issues are still persistent regardless of the device implemented therefore it's vital to acknowledge that it still plays a substantial role both historically and in modern times which only emphasises the magnitude of its importance in any project.

Though the data being stored will only include a singular image of the participant, it still carries a considerable ethical implication. This proves that the quantity of data being stored is irrelevant, it's the consistency that causes the potential issue. Encumbered under the responsibility to manage this type of data warrants the need to

inform the participants exactly how and when this type of data will be extracted. It's also crucial to articulate to those taking part the aims and objectives therefore the medium of an information sheet will be used.

For any project that involves ethical consideration, an information sheet provides in-depth details regarding the most commonly asked questions. This includes specifics about the projects mission and what is expected from the participants. It's a technique employed to offer contact information for any future questions or queries. It can also be used to relax the participants allowing them to prepare themselves for their responsibilities. A vital element was on emphasising that the data shared was going to be handled with due diligence and their anonymity will be maintained throughout the project. It's critical to elaborate on the exercises that will take place allowing for any preparation time needed. Expressing the opportunity to drop out the study at any point is also explicitly stated. The information sheet can be found in Appendix F.

It's important to make those involved in the study comfortable as working in a relaxed atmosphere will increase the chance of producing authentic results. Understanding what to expect from the project is depicted by the information sheet however accepting and agreeing to the conditions is also essential. A consent form will be given to each member of the study to reiterate their understanding of the project's requirements. It's imperative that all participants read and sign before any live testing is done. Though they will not be physically experimented on, the form demonstrates their acknowledgment into providing the necessary data for the project's progression.

Capturing and storing data sensitive information warrants the need for adequate protection. Without any additional security, hackers will easily be able to retrieve the files thus exposing the identity of those involved in the study. Implementing folder encryption such as EncryptFS will provide the defence needed to prevent any data leakages. For an added safety layer, a username and password will be required before logging into the prototype.

As the files stored are imperative to the completion of the project, the raw data and folders created will be externally backed up. This is to avoid any downtime to the overall project development and will used as an operating system restore point if any hardware or software failures occur. It's important to note that the facial recognition functions cannot be backed up as they're installed internally on the devices hard drive.

The external hard drive will also be encrypted using the same algorithm (EncryptFS). This is to prevent any potential leakages if the external HD was misplaced.

In the event that any compromising information was leaked or that the prototype was stolen, a risk mitigation plan will be executed to minimise the negative influences to both the study and the participants. The most appropriate strategy to apply would be to accept the risks involved. Risk acceptance purses the mentality of identifying the ones prominent in the study and allowing them to happen. (Herrera, 2013) In this project, implementing this plan is more effective over types such as limitation or avoidance due to it being cost effective. Mitigation using the others available will have detrimental financial effects on the budget therefore it's cheaper to simply tolerate the impacts.

To respect the anonymity of the participants, any sensitive data collected or produced will be kept confidential. Facial images displayed inside this report will be solely used to illustrate the findings and outcomes of the study. Names will also be represented by using an alias scheme. For example, '*Photo 1 shows that Participant A'*. Photos that contain any element of a participants identify will be edited thoroughly using specialist manipulation software to distort and blur information that could be used against them. Once the project has concluded, both the data collected and any results produced will be eradicated from all storage devices. This includes both the prototypes internal storage and any external hard drives used for backups. The CMD line interface will be utilised to remove all data accumulated in addition to any potential hidden files or folders created in the process.

It cannot be stressed enough that the system developed is a prototype therefore the amount of data being stored within the device will only be for testing purposes and will not be distributed after the study has concluded. It's also important to acknowledge that the study conducted is for educational purposes and the system proposed will not be implemented within a live classroom environment.

Chapter 3.4 Resources and Participation

Approaching suitable individuals to take part in the project requires thought and planning. For any project, choosing appropriate candidates is key to a successful outcome. It's vital that they possess the correct characterises and personality for an effective working relationship and contribution to the project. Therefore, the most

adequate people to contact will be fellow students from my cohort. Recruiting people from a similar course allows for related academic interests to be shared thus forming a healthy bond. If there is a lack of interest in the project, close friends and family members will be approached to fill any required vacancies. Non-vulnerable individuals will only be chosen to take part in the study.

Conducting the research in a safe and secure manner is essential for the progression of the project. This is especially the case during the data collection stage as extracting images from the participants will need to be done in a secluded environment. This is to minimise any potential unwanted identification of those involved thus maintaining total anonymity. Both data collection and testing will take place within the university grounds inside a pre-booked room. This is to provide a protected and relaxed atmosphere during the crucial stages of the study.

As discussed before, there will be 2 exercises in which the participants will be involved with. The data capture session will take around 20 minutes per individual. This is to ensure a suitable image of their face is captured for analysis. This comprises of a JPG image in which their facial features are required to be clearly visible. The live video exercise will take approximately an hour per participant or less if a group is formed. This particular experiment needs to be run multiple times in order to assess its functionality when being exposed to different environmental variables. In total, the duration of the study should last around a week. This includes the data collection stage through to the analysing the results. All participants will be informed before the project commences via the information sheet whilst dropping out the study at any time will be verbally reiterated and strongly emphasised.

One of the most considerable advantages of conducting this research is the cost effectiveness of the resources. The Raspberry Pi being built as the prototype boasts the specifications of a low budget computer including a 1.4ghz microprocessor (Heath, 2019) It is essential that the outcome of this project is exclusively a prototype system therefore utilising its versatility and easy-to-use OS makes it the perfect deliverable. Other hardware such as the wireless keyboard and USB mouse are also low-cost and can be easily obtained from any major electronics retailer.

Selecting an apt USB camera was paramount into the production of results therefore investing in a high-value peripheral that supports a 1080p or 720p HD output is vital.

The software being installed will be free meaning any development suites, IDE's, GUI's, and SQL databases created will be built with no cost required. This also applies when employing the facial recognition algorithm as a free template will be available to download from resources online. Navigating through the OS will be achieved via a 5" LCD touchscreen. This will be to demonstrate a portable and compact deliverable whilst exhibiting the features of the python GUI.

Chapter 4 – Implementation

Chapter 4.1 Front and Back-End Design

Implementation of the solution can be accomplished through testing of the proposed system. This project encompasses not only building the prototype but also the development of both front and back end architectures. The front end will contain aspects that are usable by the operator. This includes software installation but more importantly, coding of the Python GUI. The backend will allow for the student's attendance to be recorded via an SQL database. Due to limitations of hardware, the database will be designed with a simplistic approach and hosted locally on the device.

The hardware is the initial concern when planning on how to implement the solution. It was important that the specifications of the prototype are proficient enough to run the functions needed to solve the research question. The focus was building a compact system which offers maximum transportability by eliminating peripheral wires. An attempt to achieve this was by investing in a wireless keyboard which can be seen below in figure 7. You can also see the presentation of the LCD screen. This allowed for a clean finish by aiding into the portability of the system. It's also important to notice the USB battery pack powering the prototype. By creating a moveable power source eliminates the need for plugging it into the wall socket. However, this could cause an issue as providing voltage to both the motherboard and LCD screen simultaneously using a power pack battery could be insufficient.

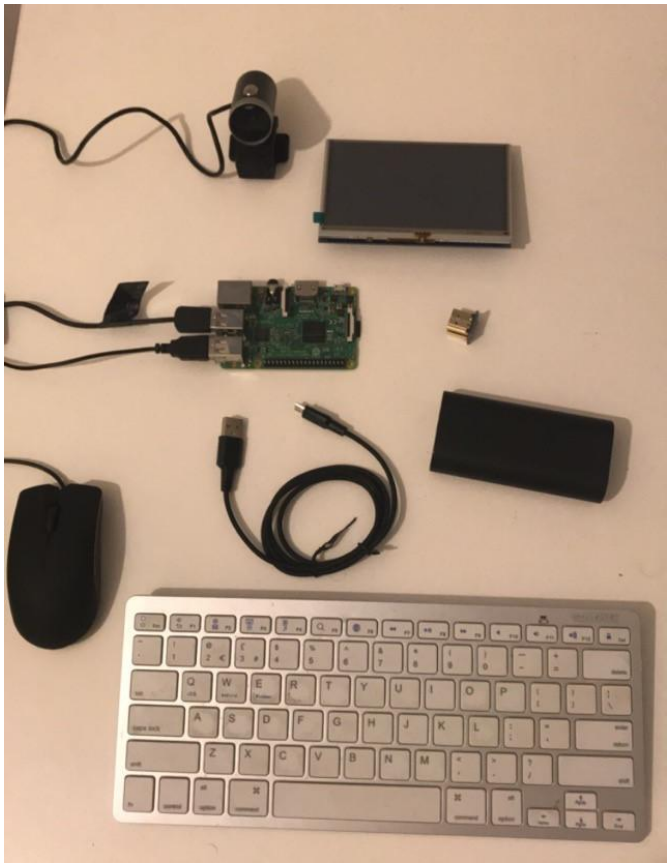


Figure 7 – *Raspberry Pi Prototype and external peripherals*

Software installations is the next consideration when designing the front end of the prototype. A Raspberry Pi traditionally is compatible with a host of Linux operating systems thus choosing to exploit the advantages of Raspbian seemed a viable option. This OS offers an easy-to-use interface for that also incorporates a range of pre-installed IDE's. The functionality behind the facial recognition is provided through the use of OpenCV. This library consists of various deep learning algorithms used to develop computer-vision oriented applications. Being open source, allows for effortless implementation within a magnitude of operating systems making it the perfect software for this project. Its architecture comprises of terminal based installations meaning only the required modules are needed. (Open Source Computer Vision, 2019)

As Python comes preinstalled on most versions of Raspbian, we only need to retrieve OpenCV's library which can be achieved by running the code below.

$ wget -O opencv_contrib.zip

https://github.com/Itseez/opencv_contrib/archive/3.0.0.zip

$ unzip opencv_contrib.zip

The above segment contains the command that downloads the required package from the online repository, GitHub. The second line unzips and installs the software directly onto the hard drive.

In order for OpenCV to function successfully, it needs to be run from inside a python virtual environment. Creating a virtual environment (VE) provides a separate work space for OpenCV to operate within. This includes segregating all modules and libraries from the standard Python 3. The advantage of this approach allows for the prevention of any potential software issues that may occur when installing new modules. However, this does also require for the correct packages to be installed within the virtual environment itself.

Before any VE can be created, a name must be given. For the sake of simplicity and purpose of the prototype, it shall be called 'cv' as shown below.

$ mkvirtualenv cv

We are required to enter the VE before any further module installations can be made. This is achieved by running the code below in which an indication will be shown that the environment has been entered successfully.

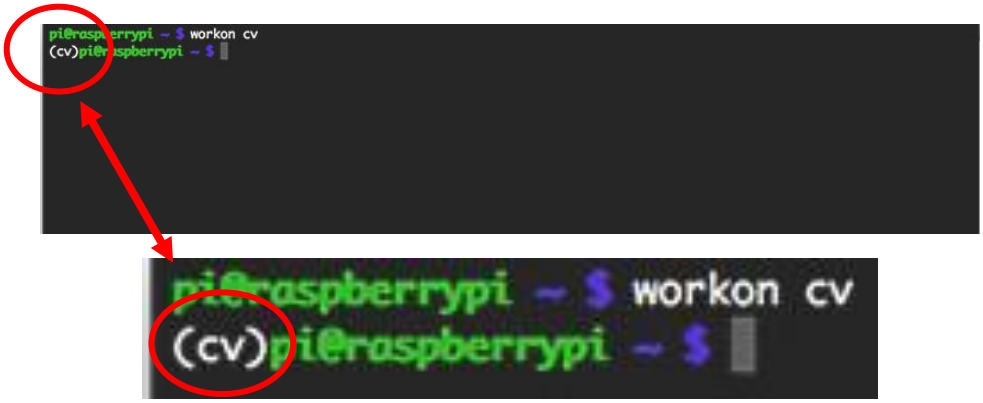


Figure 8 *– Entering the Virtual Environment*

OpenCV is the facilitator of the essential software but alone doesn't possess the ability to execute detection functions therefore, we now need to install the required facial recognition modules. The first is the dlib library in which contains the implementation for the deep metric learning algorithm. Its purpose is to build the facial embeddings from the participants images to be used for identification. The final required module is the facial recognition library. This fundamentally wraps around dlib architecture allowing it to be easily employed within the front end application development. (Rosebrock, 2018) Both are elicited from OpenCV therefore GitHub isn't needed.

$ pip install dlib

$ pip install face_recognition

To ensure successful installation and verification of the vital modules, a list can be created via the terminal. It's important to regularly check as due the amount of hardware resources needed to install these large packages, there is a potential to overheat or even crash the system during the process.

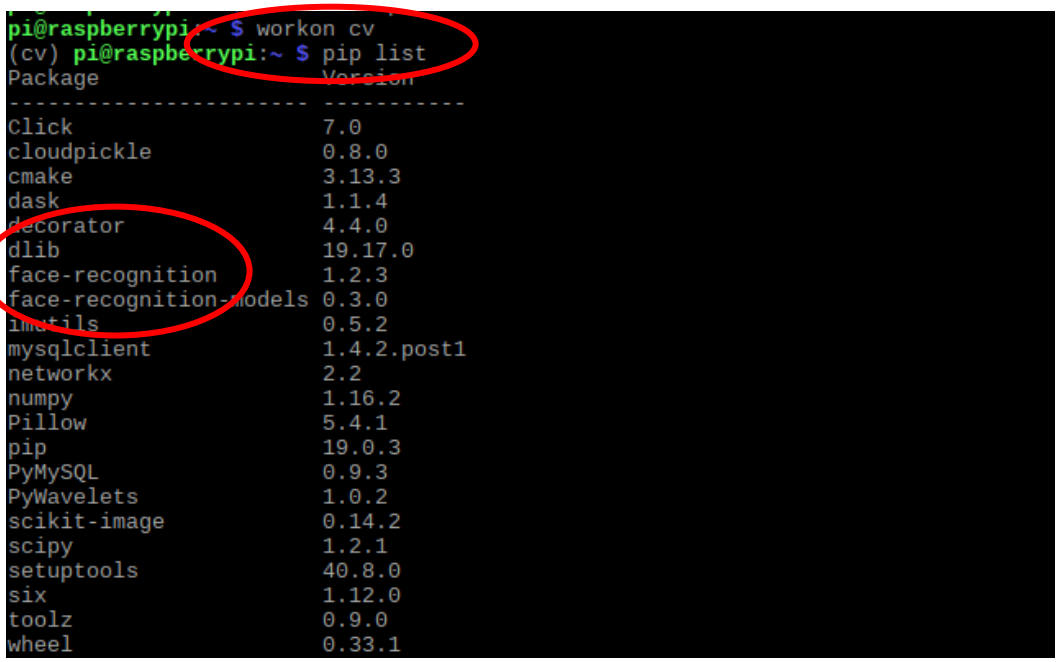


***Figure 9** - Module Check via terminal*

It's imperative to enter the VE before checking the modules as a different list will be displayed.

As all the necessary modules are now available for facial recognition, we can now move onto acquiring the python scripts to create the working architecture. The essential materials and templates are downloaded from online resources such as PyImageSerach (Rosebrock, 2018). The approach taken was to first study the scripts and ensure a thorough understanding of the existing code was made. The code could then be adapted to suit the functionality needs of the project. The template structure acts as the working environment in which any functions can be executed. Therefore, it's crucial to navigate to this directory before any testing commences.

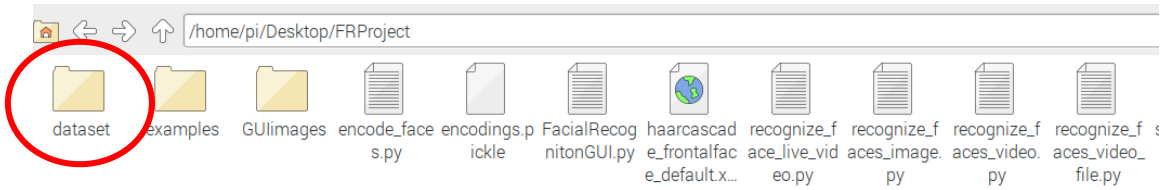


Figure 10 *– Project File Structure*

For easy access, the projects data has been stored on the desktop. The dataset folder is the file directory in which the participants facial images will be saved. These images will be used during the testing phase when being compared to other still images or during the live video exercise.

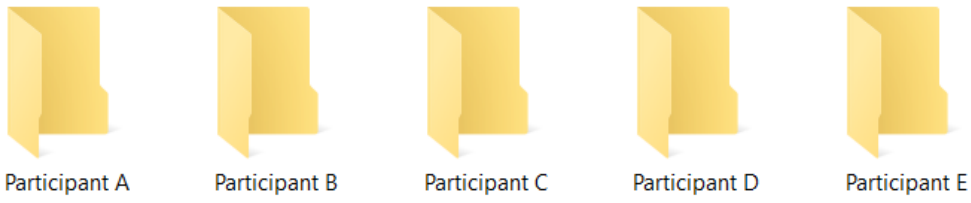


Figure 11 *– Participant Facial Image Folder*

The above image illustrates each participant of the project represented as a folder. These are subfolders of the 'dataset' folder and the names assigned will be displayed if any identifications are made during either exercise.

We now have the required resources and software installed to run facial detection commands. However, the most prominent issue involves that each can only be executed through the terminal. This includes vital functions such as entering the virtual environment, taking a photo and running the algorithm on stored images. This is a

lengthy process therefore by designing and developing a GUI will reduce the time needed to execute a command all whilst exhibiting a pragmatic interface. The python GUI is coded with a minimalistic approach as it contains 4 buttons each responsible for the key functions identified before.

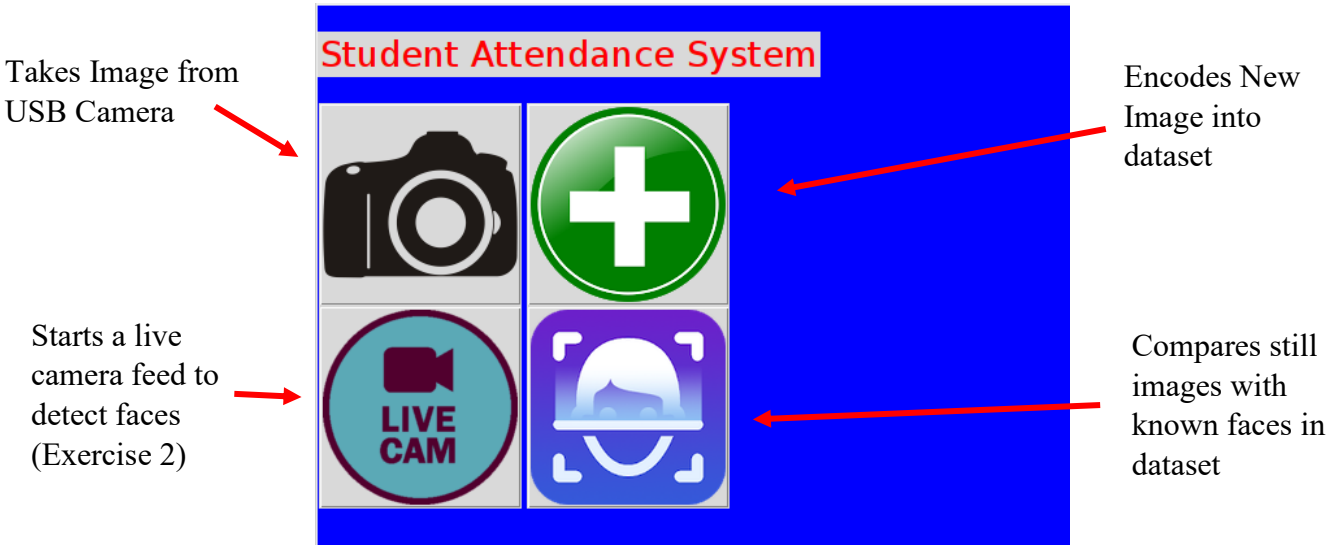


Figure 12 *– Python GUI Creation*

The interface was also designed with the consideration to incorporate the 5" LCD touchscreen. As the dimensions of the screen are 800x400, it was ideal to make the resolution of the GUI identical.

The last step is to design a backend that is capable of recording the attendance results of students. The most suitable approach was to implement a MySQL database hosted on the local machine. The focus to was to create a singular table that represented a classroom containing the participants of the study.

To ensure the successful creation of a local MySQL database, 2 essential packages are required. Initially, the Raspberry Pi needs to be converted into a SQL server providing the fundamental architecture of database establishment. Secondly, management software such as PhpMyAdmin is desirable to access and manage the local database. This can be done by running the following code segments:

$ sudo apt install mariadb-server

$ sudo apt install php-mysql

To accommodate the simplicity of the project's needs, a singular table was created consisting of 4 attributes, StudentID, FirstName, Surname and Present. Each record inserted expressed information regarding the student indicating whether they were present or absent in the class. The idea was if a particular student belonged to a class within the database and were successfully identified during the live video feed, they would be marked as present.

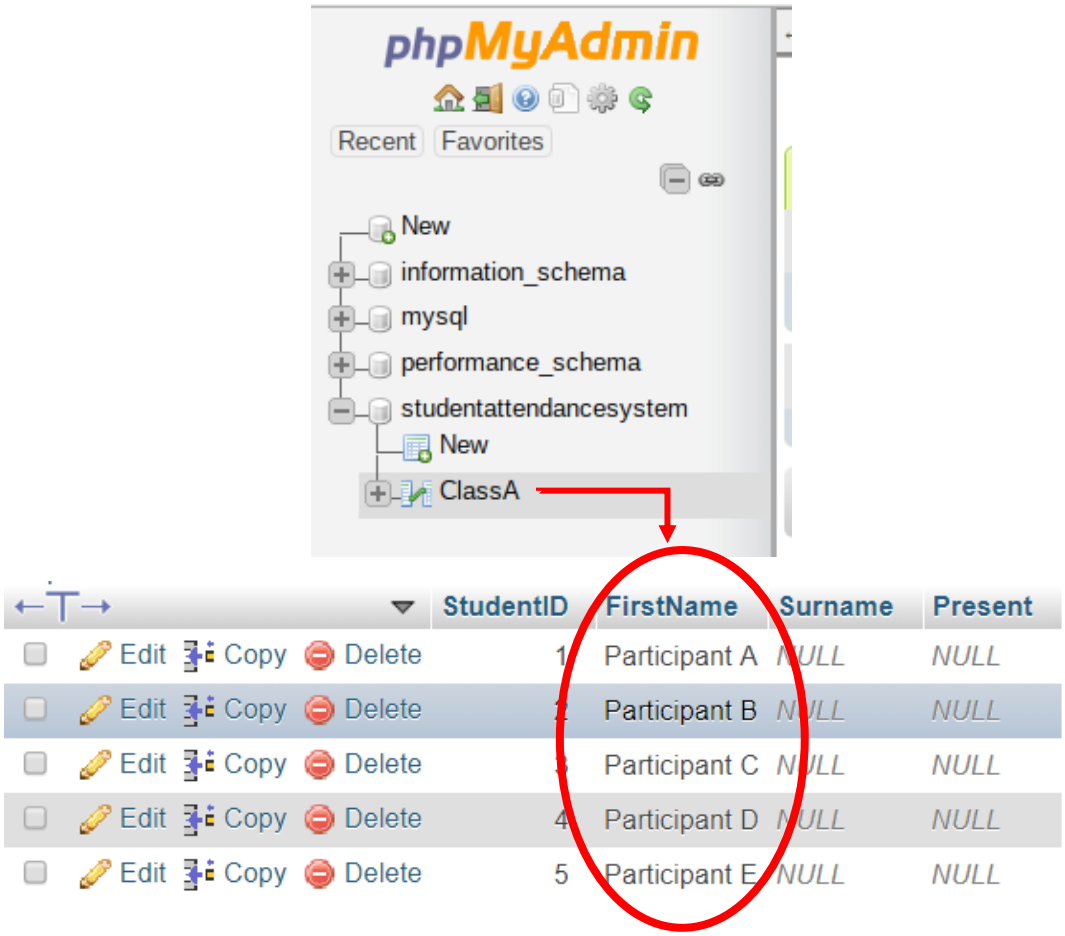


Figure 13 *– Backend Database Structure*

For testing purposes, dummy data has been inserted to represent each participant involved in the study. In a live classroom scenario multiple tables or even databases would be created to display the students full name and other respective information.

Chapter 4.2 Recognition Algorithms Comparison

Facial recognition algorithms have been historically arounds for many years. Each offer their own advantages and disadvantages over others available for implementation. The challenge is investigating an appropriate algorithm to aid into

answering the research question. We have already discovered that significant resource limitations pose a threat therefore applying an algorithm that is both capable of producing accurate results and compatible with the device's hardware is paramount.

There are a magnitude of algorithms accessible however in this comparison, we will be focusing on the 2 most frequently used in facial recognition projects. The first is known as PCA and takes a statistical approach when categorising facial images. It's a reduction algorithm that concentrates on removing the dimensions of each image providing stronger singular dataset. Its technique is to extract the most relevant information from every sample in the dataset then attempts to build a computational model known as an eigenface (Himanshi, 2015).

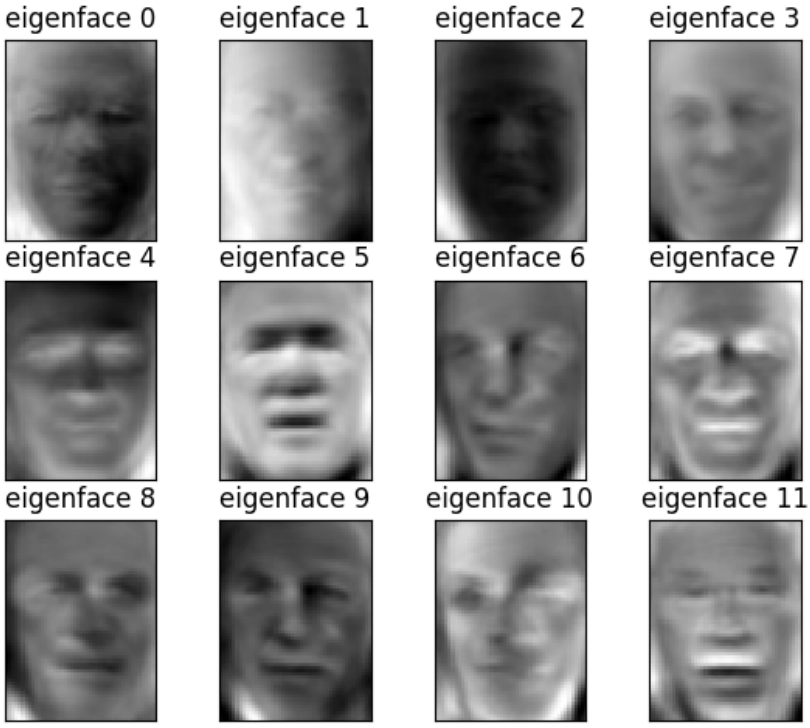


Figure 14 *– PCA/Eigenface Representation (Himanshi, 2015).*

The second being examined is the known as deep metric learning. Instead of outputting a single label, a valued vector is used to quantify the faces within the dataset. Neural networks are utilised to train the images by implementing triplets. The full process is illustrated in the figure below.

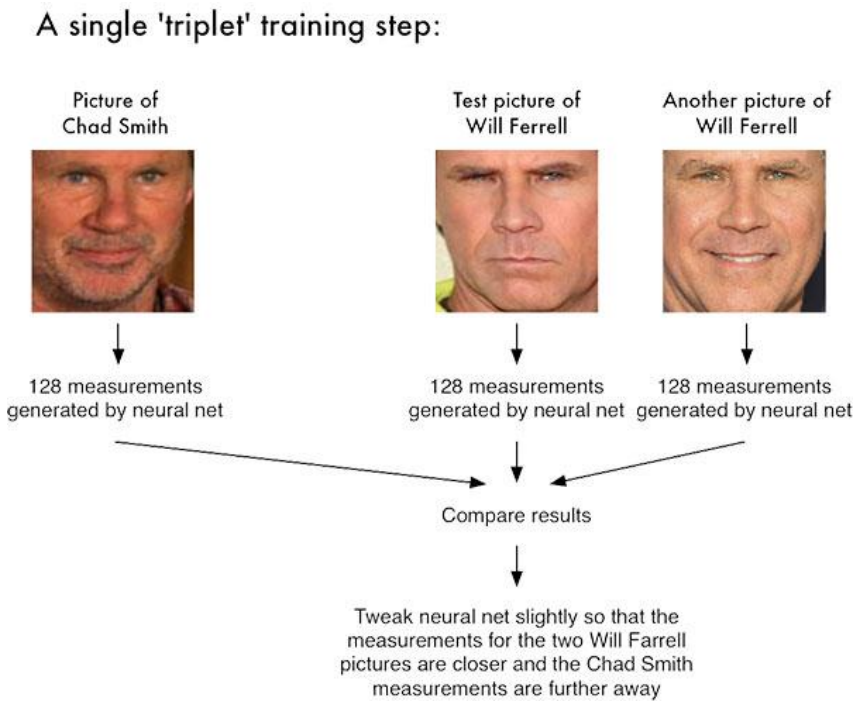


Figure 15 *– Deep Metric Learning Triplets Representation (Rosebrock, 2018)*

In the above example, we can see 3 images that have been encoded within the neural network. Two of the faces are of the same person whereas the third differs however they each share similar facial features. The network quantifies each image building a 128-dimentional embedding for later identification. Weights of the network are altered so that the 2 similar 128-d measurements will be closer to create the detection separation from the third image. (Rosebrock, 2018)

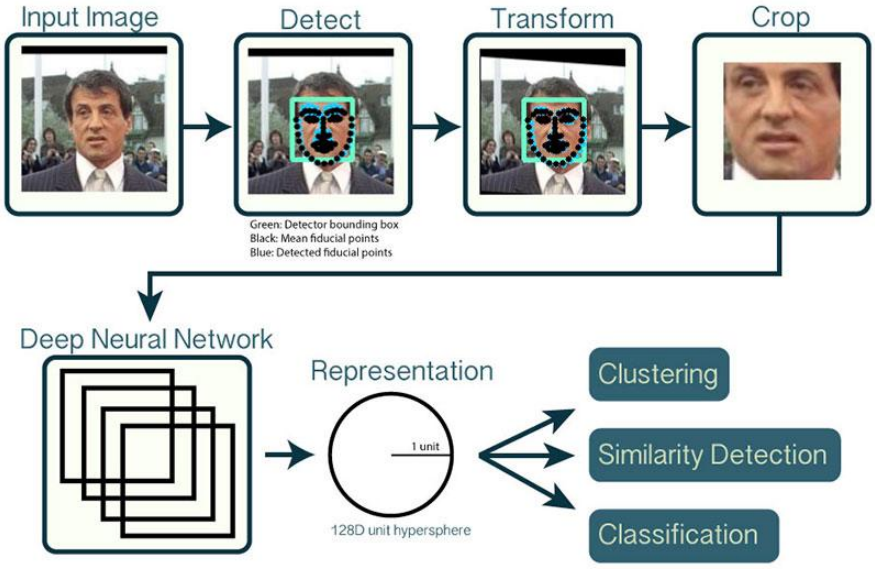


Figure 16 *– OpenCV's Facial Recognition process*

One of the most substantial disadvantages of implementing PCA for this particular project is that its process involves scanning the entirety of the faces surface area. Though this approach successfully quantifies each image in the dataset, it subsequently will require an extensive amount of hardware assets. Due to the lack of resources the prototype has available to distribute, will inevitably produce less accurate results. On the other hand, applying the deep metric learning algorithm will allow for a faster and more efficient facial detection rating as it takes a feature extraction approach rather than quantifying the whole face E.g. pose, expressions etc. Providing more images that could be encoded within the dataset elicits higher accuracy of results (Hartnett, 2015) however due to storage limitations of the prototype, this may be an issue if the capacity wasn't increased.

Chapter 4.3 Testing Environments

Scheduling the exercises was planned around the availability of the participants allowing them to take part whenever suited them best. It's vital that those who are taking part had returned their consent form and agreeing to full participation. (Appendix G) A verbal confirmation was also given into their acknowledgement of the information sheets substance and confirmed they had understood the details provided. These essential documents must be retuned before any data collection or testing can take place. It's also a method provide clarity into the projects overall rationale and articulating its objectives.

For optimum results, the approach taken was to either set up a group exercise or to test the prototype on each participant. Testing on a large group would allow for extensive analysis of the algorithm employed. Authenticity in the results could also be obtained as with more members, a closer representation of a classroom could have been made. However due to the clash of personal commitments, each took part in the study individually. During a particular instance, 2 members were present for the exercises therefore photos of this session will be used to illustrate the methodology and results.

As identified before, the location of the study was conducted inside a pre-booked room within the university grounds to provide a safe and secure atmosphere. Exercise 1 was the data collection phase in which involved taking images of the participants faces. The photos themselves had to achieve a certain level of clarity to be used in the study. This included a clear visualisation of the main features of their face such as the eyes,

mouth and nose. Failing to capture an image of such quality will result in poor detection rates thus providing an inaccurate demonstration of the prototypes performance. The optimum setting was to create a well-lit room using as much natural lighting at possible. The angle of the camera had to be correct and centred on the face to cover a suitable surface area. Distance was also key as taking the image too close will result in a blurred and unfocused picture.

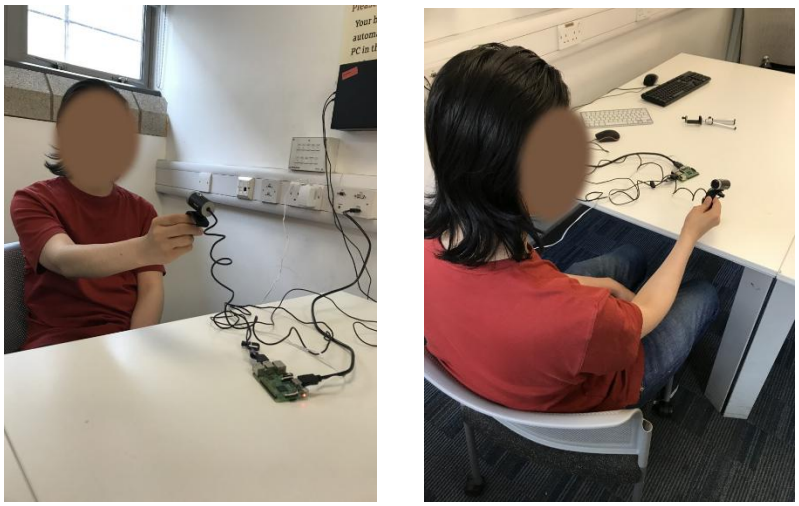


Figure 17 *– Photo Extraction Exercise*

Recognising the best setting to take a practical image was strenuous as there were multiple variables that effected the picture quality. The main concern was the reliability of the camera itself. Though it supported the output of 720p HD images, the prototype system didn't possess either the software or hardware capabilities to utilise the properties of this feature.

Capturing an image using the terminal for each participant caused a major interference on the project's progression. Typically, when executing the code, a photo is taken and saved within the root directory of the current user logged in. This was sufficient during the initial hardware testing phase however, for this exercise to be a success, the GUI needed to be exploited. Using the GUI developed allowed for a smooth and systematic process through the entire study as each button ran the required code in a step by step fashion. The first stage was using the GUI to name the participant which was taking part in the session.

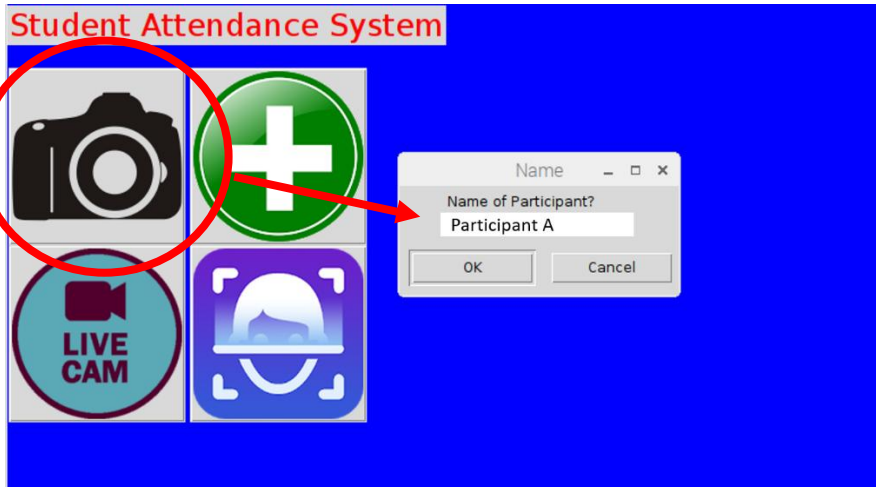


Figure 18 *– Participant Creation*

Each user conceived through the GUI entails the creation of a directory within the dataset folder labelled with their name. If their data already exists, a message box will show. Once the operator is satisfied with the name inserted, an image of the face is also taken simultaneously during the same process. The photo is then transferred to the new participant dataset folder. This image can now be used as their baseline photo for comparative purposes when either detecting from a still image or during the live video feed.

Before any biometric identification can occur, the facial image needs to be encoded. This is the essential stage that applies the principles of deep learning metrics to quantify the embeddings of the subject's picture. By completing this integral step will notify the system that a new user has been added to the dataset.

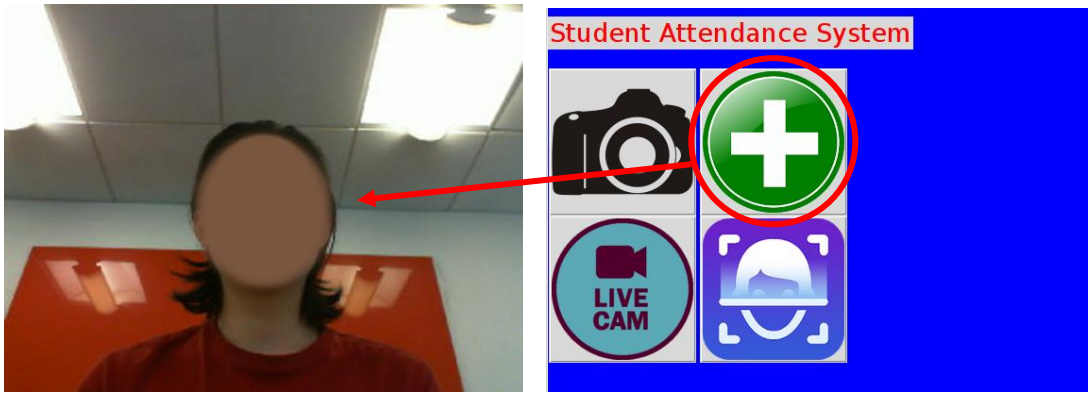
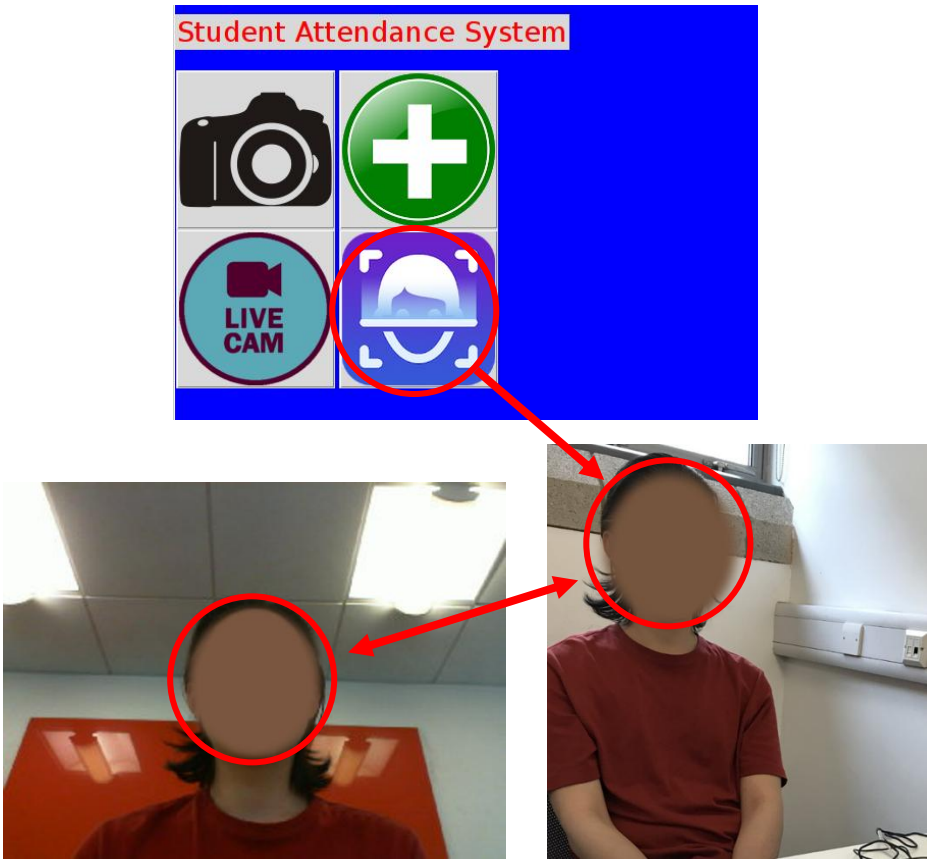


Figure 19 *– Encoding of new participant*

To test the successful encoding of the image, we can use a function that allows us to compare it with an alternative photo of the participant. Doing this will measure the quality of the initial image taken providing the opportunity to retake if necessary. A

gauge can also be obtained into the accuracy of the facial recognition algorithm applied before commencing the live video exercise.



Encoded Image

Comparison image for accuracy indication

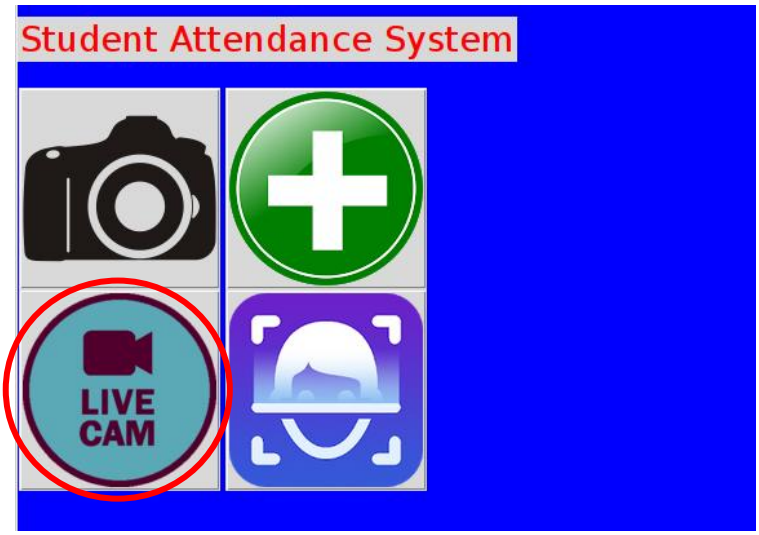
Figure 20 *–Encoded image vs alternative image comparison*

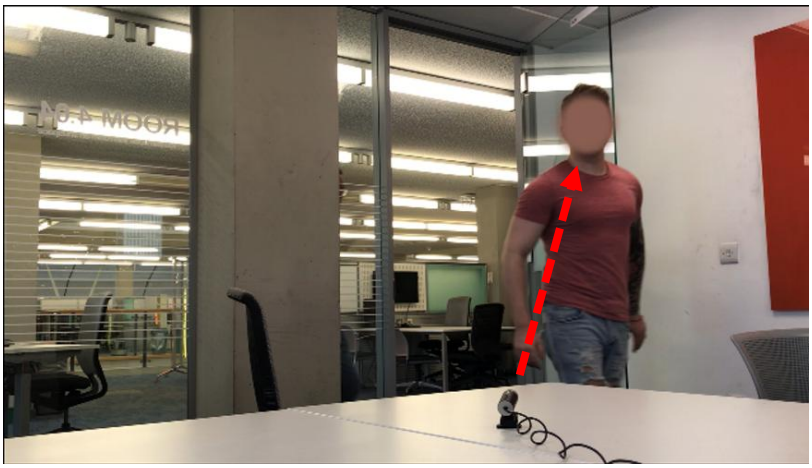
Now the participants images have ben encoded into the dataset, we are able to proceed onto the second exercise of the study. This experiment sees the participants simulate the beginning of a university lesson in which they enter the classroom to be scanned by the camera. Assessments will be made to monitor the prototypes performance in successfully identifying each participant face as they arrive. It was essential to position the camera in an area that can both acquire a clear view of the faces but also demonstrate its ability to scan at a distance.

To start the exercise, the GUI is used as it contains a button that binds the command to start a live feed that runs constantly until the camera is closed by the operator. It will be interesting to observe the consistency in accuracy made by the prototype as

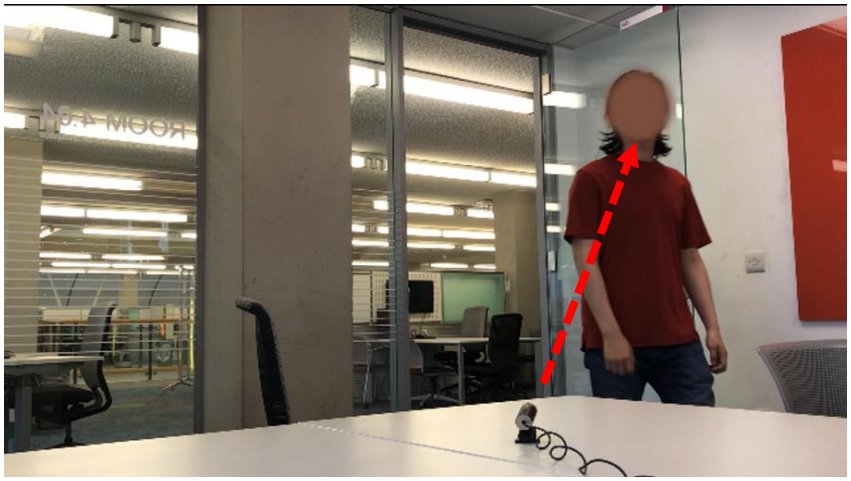
executing a live video will allocate a considerable amount of resources when being compare to a still image comparison. An issue that may occur is the loss in frame rate. Though using a smaller screen to view the live video will be used, the system still relies on the prototypes hardware to maintain a high FPS. This could be another factor that effects the end results.

Acquiring a wide range of results will be beneficial into gaining a deeper understanding of the prototypes performance therefore, this exercise involves 3 smaller experiments. The first sees each participant to enter the room independently then as a group. Applying this technique will allow the system to demonstrate its efficiency when detecting 2 faces in the same frame.

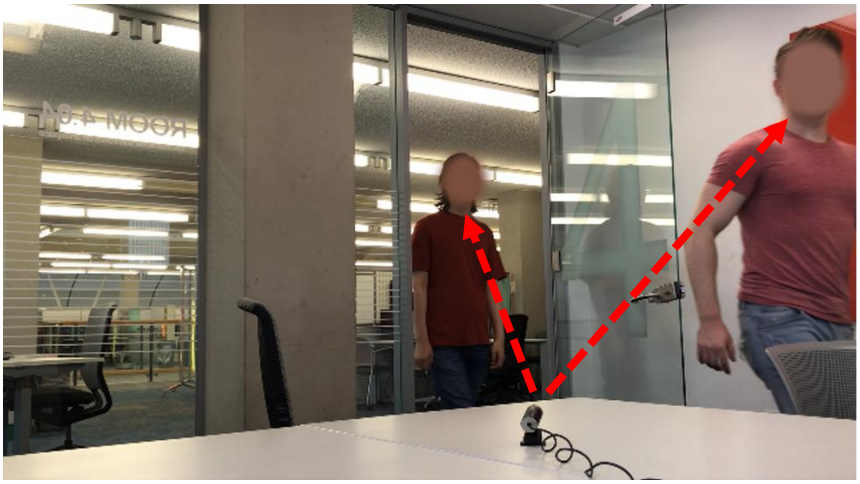




Participant A



Participant B

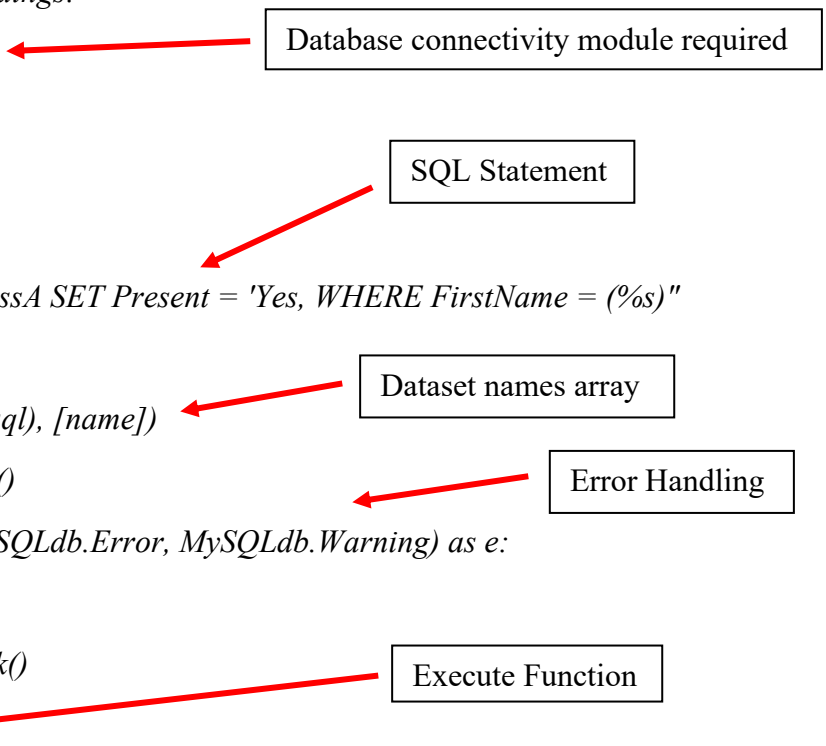


Participant A + B

Figure 21 *– Live Video Feed Exercise*

Implementing a method in which to capture the attendance of the students is vital into discovering a potential research question answer. Using the MySQL database design as mentioned before is the perfect solution to record the student's overall presence. Adapting the already existing template to perform SQL functionality is a challenge therefore understanding the initial scripts was crucial before adding the required database connectivity code as shown below.

```
for encoding in encodings:
import MySQLdb
global c
global db
def insert_to_db()
sql = "UPDATE ClassA SET Present = 'Yes, WHERE FirstName = (%s)"
        try:
            c.execute(sql), [name])
            db.commit()
        except (MySQLdb.Error, MySQLdb.Warning) as e:
            print (e)
            db.rollback()
    def main():
insert_to_db()
```



It's important to notice the line '*for encoding in encodings:'* as this allows for the above function to run every instance a known face has been detected during the live video exercise. Each time the code is fired, the system is comparing with the local database to identify if the participant exists within the table 'ClassA'. If a match is discovered, the software runs the SQL statement to update the database record changing the students 'present' attribute field to 'Yes'

Chapter 5 – Evaluation of Results

Chapter 5.1 Expectations Vs Results

As the project began to expose more knowledge in this field of study, it starts to mould the interpretation into the consistency of the outcomes. Examining previous literature has provided a sense of clarity when approaching this topic from both a technical and ethical point of view. These crucial aspects of any technology-based project has formed a benchmark into what to expect and also how to use the results to improve the quality of research for future endeavours.

From the beginning, data capturing seemed to be effortless from a technical standpoint however this wasn't this case as capturing a suitable image of the participants was a

constant challenge throughout the exercises. Taking into consideration multiple environmental factors such as the angel and lighting will make a considerable difference regarding the picture suitability. Relying on the 720p HD camera was easy as it was expected to take quality images irrespective of the system it's been implemented on. This proved difficult to believe as it took several attempts to capture a clear image of the participants face. This had a severe effect when trying to carry out the rest of the study as poor image quality will produce poor detection results.

Once an adequate image had been captured, the initial test was to ensure the prototype could correctly identify the participant on both the same and an alternate image. This confirmed that the systems detection capabilities were working.

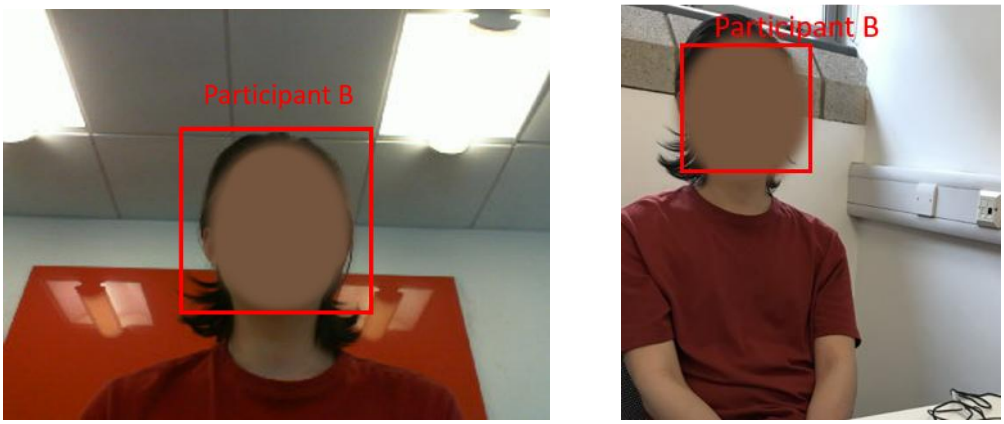


Figure 22 *– Still images comparison result*

As a result of exercise 1, the detection has correctly expressed a full match in both images as the system has labelled them Participant B. The image on the left was captured initially and encoded within the dataset. This alone should pose no issue regarding the identification as it's being compared against itself. The image on the right was taken during the same exercise however from a slightly different angle. There were no detection issues presented when a comparison was made.

Establishing that the system possesses the ability to detect faces accurately, assessing its performance during the live video exercise ensued. As discussed previously, this experiment involved running 3 activities seeing each participant entering the room independently and then together. In theory, the system is expected to perform equally

regardless of the quantity of members in the same camera frame. The hardware should be efficient enough to recognise multiple individuals simultaneously.

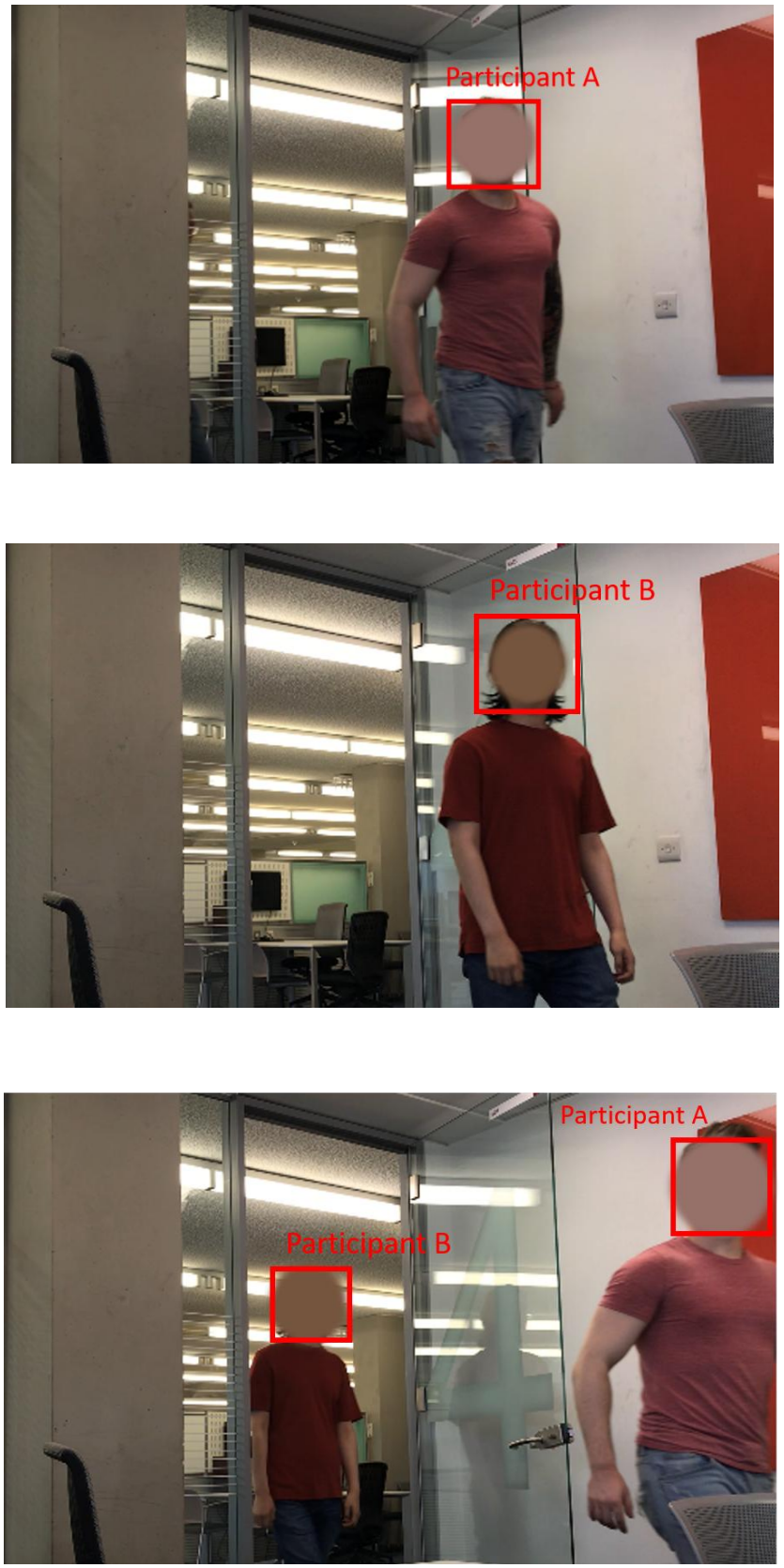


Figure 23 *– Live Video Exercise Results*

The results to this exercise were astonishing as for the majority of the experiment the prototype maintained a successful detection rate. Figure 23 illustrates this as when the participants both entered separately and as a group, the system performed remarkably. This proves that the hardware contains the processing power to sustain an accurate reading when recognising multiple faces which goes against expectations. However, on a few occasions, issues did arise when trying to consistently retain the correct identification.

The most common issue was mistaking the identity of the participants for others involved in the study. For example, in a few instances, the name displayed above the face was incorrect. Once the error did occur, it would only persist for a few moments then revert to recognising the individual correctly. The main cause could potentially include environmental effects such as poor lighting or dust blocking the cameras view. Another error could involve a temporary software glitch due to insufficient resources or a slight dip in processing power. An example of the issue is shown below.

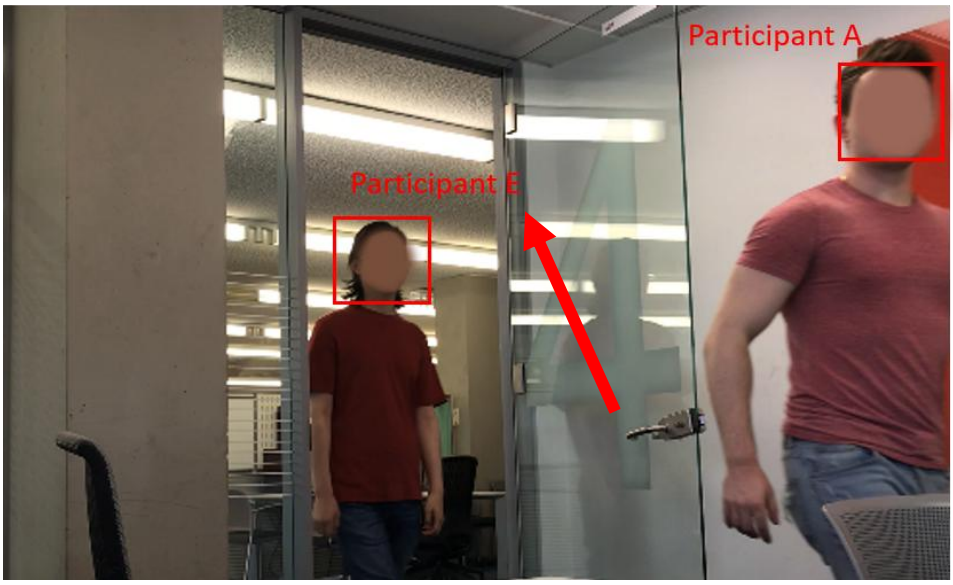


Figure 24 *– Incorrect Identification of Participant*

Another reoccurring error was the recognition of an object in the room as a person. It's arduous to define exactly what caused this issue without extensive testing however through observations, the most logical reasoning would be due to some items or wall decorations appearing to resemble a face. Another possible motive could be issues with the hardware itself. It's not uncommon that the system suffers from glitches

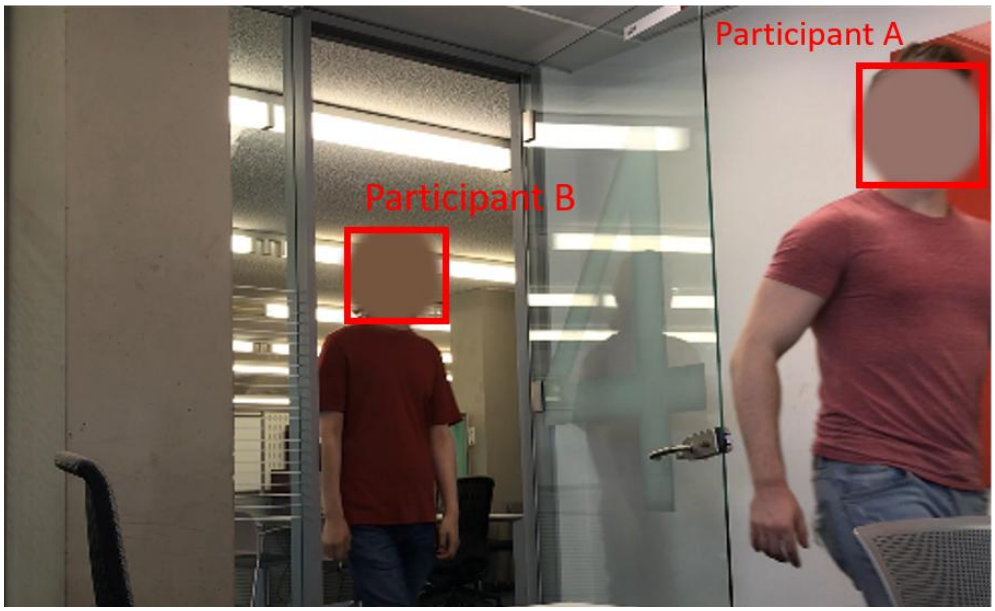
causing the algorithm to fire unexpectedly triggering spontaneous recognition of inanimate objects. The issue can be seen below.

Figure 25 *– Detecting faces on surroundings*

To ensure the students attendance was successfully recorded during the exercise, the SQL database was utilised. Implementing the functionality expected to be straight forward as there are limited variables that can affect the outcomes of tested code. Developing the database itself took minimal effort however injecting the solution into an already complex python application was a challenge. The biggest issue was ensuring the database code was inserted correctly inside the script to establish the connectivity.

Writing the code was effortless as existing experience of SQL scripting was applied however testing the performance was problematic. This was due to the fact that the prototypes preliminary features had to function as normal, this included successfully detecting known faces with no errors occurring in the process. After both the facial recognition was operating as expected and the SQL syntax correctly written, recording student's attendance could ensue as demonstrated below.

	StudentID	FirstName	Surname	Present
Edit Copy Delete	1	Participant A	NULL	NULL
Edit Copy Delete	2	Participant B	NULL	NULL
Edit Copy Delete	3	Participant C	NULL	NULL
Edit Copy Delete	4	Participant D	NULL	NULL
Edit Copy Delete	5	Participant E	NULL	NULL



	StudentID	FirstName	Surname	Present
Edit Copy Delete	1	Participant A	NULL	Yes
Edit Copy Delete	2	Participant B	NULL	Yes
Edit Copy Delete	3	Participant C	NULL	NULL
Edit Copy Delete	4	Participant D	NULL	NULL
Edit Copy Delete	5	Participant E	NULL	NULL

Figure 26 – *Recording Student Attendance Results*

Chapter 5.2 Performance Analysis

It's vital to reflect into what contributed to the results produced by the prototype. These areas need to be thoroughly analysed and measured to ensure the objectives can be met and the research question answered. It can also be a method in which to donate knowledge in this field by determining what elements worked efficiently compared to those that didn't. Discovering new information provides the opportunity to show others or reveal alternative approaches to this type of project in the future.

Stages such as taking an image of the participant will contribute to the overall accuracy of the results generated. This is because the photo captured will be used to identify each member during the still image comparison and live video exercises. That being said, the photo needs to be of satisfactory quality to confirm that the rest of the study can be achievable. Once a suitable picture is extracted, it can then be used to represent the participant for the longevity of the study.

During the data collection phase, several attempts were made to ensure an appropriate image could be used to encode within the dataset. For the majority of the study, practical images were extracted however on some occasions, it proved that capturing a useful photo was challenging. The performance of the camera was at fault as this was the tool applied to both elicit still images and conduct the live video sessions. The issue was regarding the hardware and software used to manage the cameras capabilities as discussed before.

Most PC's have inbuilt camera software and drivers that allows for automatic camera optimisation meaning that once it's installed, the system is intuitive enough to adapt the settings allowing for clear images to be taken. This can also affect the CPU's performance as the more resources available, the quicker the camera can adapt to certain environments. The location in which the study took place didn't contain any harsh conditions that would affect the camera therefore taking an image of the participants should have been effortless. On multiple occasions the images taken was either too dark, distorted or even displayed excess brightness rendering it useless for testing purposes.

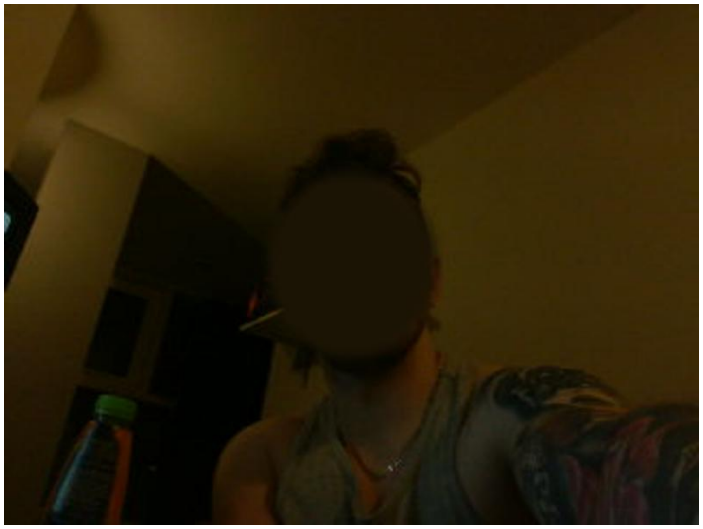


Figure 27 *– Dark Image Capture*

Figure 27 articulates though the image taken was in well illuminated environment, due to the lack of resources distributed by the prototype, the resulting image expresses as lack in brightness. Similarly, another issue involves pictures to contain a horizontal and vertical border across the edges. The cause of this would also be associated with the deficiency of processing power. In many instances, these borders would appear over the face which subsequently produced encoding errors.

Figure 28 *– Image Border Issue*

The accuracy of the facial recognition software proved to be a success throughout the project. During both exercises, the consistency of detecting the participants faces remained steady. Contributing to the overall accomplishment was two key elements such as the efficiency of the algorithm but also the quality of the initial dataset image.

The algorithm provided a neural network capable of accurately reading and quantifying facial images in an effective manner. It's low level processing needs made this approach ideal over other algorithms such as PCA/LDA. The other vital aspect is the usability of the dataset image being applied for comparative purposes. As identified previously, due to hardware limitations, capturing of a poor image is likely therefore it was imperative that a clear photo was established before either exercise commenced.

The distance in which the camera could detect was also an interesting observation as in a live classroom environment, the positioning of the camera would be unknown. Due to the spatial restrictions of the room where the study took place, testing this element was impossible. However, judging by the space accessible, this indicated that the range in which the camera could detect was adequate for any medium – large classroom.

Conducting this type of research on a microprocessor will have its limitations especially when implementing facial recognition technology. When testing the system on a still image, the prototype is able to correctly identify each participant with ease as it has the time and resources to do so. However, during the live video sessions, the system is under a considerable amount of processing stress as not only does it need to run the algorithm, it also is required to maintain a consistent live video stream which in itself consumes a substantial amount of RAM. Due to the lack of distributed resources, it was clear that the frame rate suffered dramatically.

When the feed started, the frame rate sustained 60fps however when a participant's face was detected, it dropped significantly to approximately 4fps. To counter this issue, steps were taken to lower the resolution of the screen thus reducing the load on the processor. This approach only increased the frame rate by around 5 which wasn't enough to create a more latency free video stream.

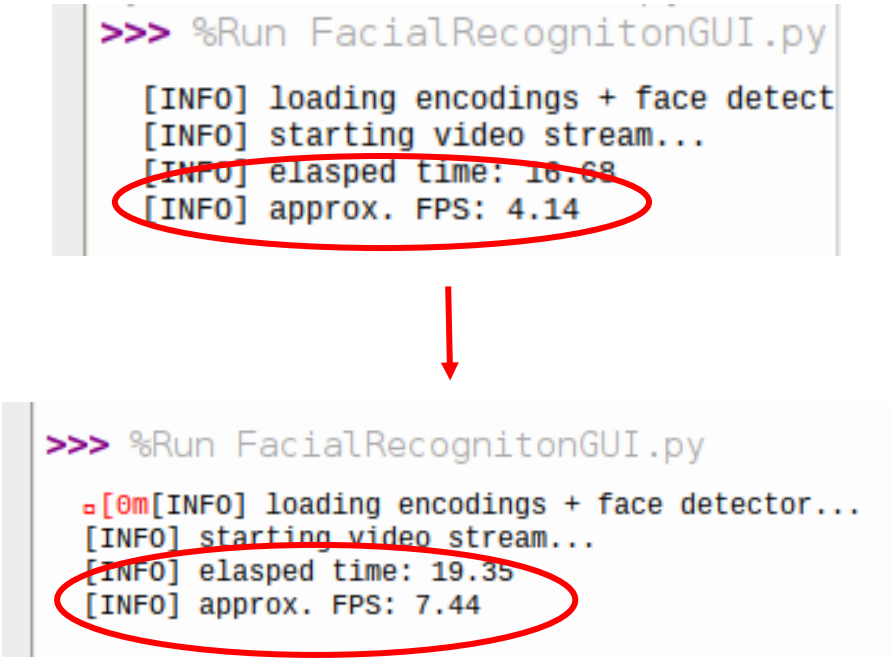


Figure 29 *– Increased FPS due to lower screen resolution*

Chapter 6 – Discussion

Chapter 6.1 Research Question Re-evaluated

Reflecting on the project conjures the thought into why it was conducted initially. It's easy to get lost in the madness of producing results and designing solutions where you begin to lose sight behind the reasoning into achieving the aims and objectives. The rationale of this paper is to identify an answer to the research question expressed in the introduction chapter but also to dissect and discuss the approach into answering it. When eliciting the question at the beginning, it seemed like an appropriate topic to study due to the uniqueness of the prototype system proposed. However, it was clear that during the study, the research question was too open and ambiguous.

It became apparent that throughout the literature review, systems that pursue the same motive already exist. Though not exactly identical, there are similar that have an exclusive approach on answering the research question stated. It was vital that I took from each paper the methodology applied and acknowledge the advantages of the respective solutions. As previous papers were discovered, the interpretation of the research question changed thus forcing it to be re-evaluated. It seemed that it needed to be narrower and address a more specific problem.

During the project's progression, it was evident that the research question started to become obsolete. It's not uncommon that as more information is accumulated, different ideas begin to be form in the mind which subsequently alternates the path in which the study follows. This was particularly the case as the implementation stage was heavily focused on hardware and software testing. Spending a considerable amount of time on this aspect changed the dynamics of the project. Whereas the focus was on initially identifying a method in which to accurately record student attendance then inadvertently reformed into a study revolving around facial recognition algorithm implementation on a Raspberry Pi.

Key areas such as the literature review has unveiled how the research question needs to be more narrow. The question defined exudes a lack in clarity regarding the motive driving the project. It seems that the concepts of Raspberry Pi microprocessors, facial recognition technology and student attendance systems have each been covered individually rather that collaboratively therefore it feels that the experiments conducted do not contribute to a specific area of research but instead provide a generic overview. If the research question was tackled with a more explicit approach, then specialised testing could be planned thus providing more constructive results in this particular field of study. Alternative questions such as how to utilise a Raspberry Pi to capture student attendance? Can facial recognition technology accurately record student attendance? could be asked.

Has the research question been answered? Yes. With traditional systems still claiming dominance in the majority of higher education institutions, it seems that the proposed prototype is able to supersede current techniques and capture attendance in an efficient manner. A considerable amount of issues with the traditional systems have previously been identified therefore the implementation of the proposed prototype will rendering existing methods obsolete thus providing a more accurate process into recording student attendance. The results produced from this study has proved that not only this type of technology can be implemented within an education environment, the resources needed to create a successful working system are minimal. Even when employed on a microprocessor, the consistency of facial detection remains high making this the ideal solution when solving the research problem.

Both historically and in modern times, ethical consideration will always pose a significant issue when applying this type of technology in any device. We have seen privacy is one of the most important concerns when in public therefore implementing the prototype in a school will require an equal amount of ethical professionalism. Existing systems such as SMART CCTV is enforced solely to prevent crime and terrorist attacks however it seems that the trade-off between personal privacy and security will always be a controversial topic.

To minimise the issues caused by this type of technology, it seems that the best strategy is making the individual fully aware that they are being monitored and recorded. If the person provides their consent and has the opportunity opt out of their data being logged then this could create a basis to utilise this type of technology both in a public space and within educational institutions. If they do provide their consent, then it's also pragmatic to express any security measures in place that protect their information and also the intent in which their data will be used.

Chapter 6.2 Limitations

Every project will contain issues that prevent it from being a success. There is no ideal venture that builds a system that has the ability to generate the best results possible. The aspects stopping any project from achieving the optimum outcomes are its limitations. Using a Raspberry Pi in any project will induce concerns regarding its hardware capabilities. Though it contains a microprocessor capable of running simple applications, applying the use of facial recognition software initially poses a risk to the projects resource stability. The main limitation was the extent of the processor power. As the device boasts a 1.4mhz speed, it should be proficient enough to run all the required functions

It was clear that during the live video session, the frame rate dropped significantly once a participant had entered the frame. The main reason behind this was insufficient processor and RAM resources which caused the system to occasional freeze. Though it did not affect the facial detection rate or recording of the attendance, implementing an unstable prototype would inevitably cause hardware failures in the future. An attempt was made to rectify this issue by reducing the resolution of the python interface. However, this only increased the FPS slightly which still prevented a smooth visual display.

The next most substantial limitation involved the accuracy of the software. More specifically, the number of occurrences that the facial recognition either produced an error or incorrectly identified a member of the study. Though the prototypes detection rates were high, it still possessed the ability to display the wrong identity. For the testing purposes, allowing this type of issue to occur provided a deeper understanding into combining microprocessors with biometric technology however implementing this system in a live classroom would cause an array of attendance issues. This again stems from the lack of hardware resources available.

Other issues that prevented the project from reaching its full potential include the abundance of wires present when operating the system. One of the main focus points was building a professional artefact that could have the potential into becoming portable. Due to the dimensions of the motherboard, each USB peripheral connected would easily become tangled and could lead to hardware damage. An attempt to employ wireless equipment was made as a mobile keyboard was utilised for the study. If a wireless mouse or even camera was incorporated, this would have reduced the amount of wires thus creating a more portable prototype.

The last limitation of the project would include the amount of storage available on the device. Though it didn't pose an immediate threat, if an extensive amount of data was collected, the memory card would imminently run out of capacity. As JPG images are the prime data being captured, investing in a large memory card was sufficient for this study. Not only are the images being stored, the SQL database is hosted locally. This will also take up a substantial amount of space especially if the database grew in volume. For future expansion, the SQL database will need to be hosted externally on its own server or seek to increase the internal storage size.

Chapter 7 Conclusions and Further Work

Concluding a project will entail reflection of the processes taken into building the artefact. This chapter can also be utilised to distinguish whether the aims and objectives have been satisfied. Commencing a project involving any concept of information technology creates the ideal opportunity to exhibit the skills accumulated over the course. The passion for Raspberry Pi's and appetite into diving deeper into the world of facial recognition fused together perfectly allowing for the manufacturing of a successfully working deliverable.

The construction of a successfully working system in alliance with the depicted methodologies and implementation strategies found in the project proposal has been made. The concluding iteration has considerably adapted when being compared to the first version developed. This is because discovering that similar systems exist and the ethical implications caused by this type of technology forced the approach to completely change. This expresses an aptitude into acknowledging modern technologies and society considerations to then be implemented in a way that improves the proposed solution.

Many additional features could have been included in which would have enhanced the functionality whereas some were for an aesthetic contribution. Through further time and resources, integrating these tools would improve the prototype providing it with more proficiencies into producing increased result accuracy. The most substantial modification for future development of the system would be to apply the same software architecture but on a desktop computer. By utilising the hardware capabilities of a PC such as an Intel or AMD processor will provide substantially more resources during all aspects of the project. There will be an immediate difference in improvements regarding both the live video session's FPS and detection results. More resources can be distributed to the core functions of the system such as running the facial recognition algorithm which inevitably will maintain higher identification accuracy and consistency.

To gain a better indication into the full accuracy of the deep learning algorithm applied would be test the prototype on a wider audience. A limitation of the project was the lack of participants gathered during the implantation phase. Conducting a study on fewer individuals sets a boundary into the validity of research and findings discovered. The focus was primarily on observing the accuracy of the algorithm applied therefore testing it on a maximum of 5 participants isn't enough to fully test is capabilities. The more members involved in the study would provide a broader set of results which then could be recorded and analysed for future improvements.

As we have diagnosed already, there are a magnitude of various facial recognition algorithms each boasting their own advantages over others on the market. The challenge is selecting one that is most suitable for the system being developed. In this instance, the deep metric learning algorithm applied was most appropriate over the

PCA/LDA however that was only one comparison being investigated. Due to the number of others available, its ambitious to ascertain which one is optimal when implementing for this particular prototype. If more time and resources were accessible, an in-depth analysis could be undertaken to identify which algorithms could produce improved results. A statistical approach could also be taken as by recording the detection rate as a percentage would make it easier to compare the effectiveness of the chosen algorithm to others.

I hope with the research provided in this paper, a deeper understanding of facial recognition technology can be ascertained for pragmatic uses and for the protection of society. It's clear that biometric identification is becoming increasingly integrated into a number of devices therefore I'm excited to see what this type of technology has to offer in the future.

References

Epstein, J. (2010). Present and Accounted for: Improving Student Attendance Through Family and Community Involvement. The Journal Of Educational Research. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/

Jha, A. (2015). Class Room Attendance System Using Facial Recognition System, Vol. 2(Issue 3). Retrieved from https://s3.amazonaws.com/academia.edu.documents

Alterman, A. (2003). "A piece of yourself": Ethical issues in biometric identification. Ethics And Information Technology. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.455.4226&rep=rep1&type=pdf

Kar, N. (2012). Study of Implementing Automated Attendance System Using Face Recognition Technique. Computer And Communication Engineering. Retrieved from https://www.researchgate.net/profile/Nirmalya_Kar/publication/

Krishnan, M. (2015). Implementation of Automated Attendance System using Face Recognition, Volume 6(Issue 3). Retrieved from https://pdfs.semanticscholar.org/61ab/3f2d787d30538af9cbe63a5d4b4cb62676a8.pdf

Lukas, S. (2016). Student attendance system in classroom using face recognition technique. Retrieved from https://ieeexplore.ieee.org/abstract/document/7763360

Brey, P. (2014). Ethical aspects of facial recognition systems in public places. Information, Communication And Ethics In Society. Retrieved from https://www.emerald.com/insight/content/doi/10.1108/14779960480000246/full/html

Nissenbaum, H. (2010) Protecting Privacy in an information age: The problem of privacy in public. Law and Philosophy, 17: 559–596.

Cadwalladr, C. (2018). The Cambridge Analytica Files. Retrieved 28 August 2019, from https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

Bowyer. K. (2004). Face Recognition Technology: Security versus Privacy. Retrieved from https://ieeexplore.ieee.org/abstract/document/1273467

Fuzail, M. (2014). Face Detection System for Attendance of Class' Students, VOL. 5(NO. 4). Retrieved from http://www.ijmse.org/Volume5/Issue4/paper2.pdf

Sajid, M. (2014). A conceptual model for automated attendance marking system using facial recognition. Retrieved from https://ieeexplore.ieee.org/abstract/document/6991407

Rhesa, A. (2014). Implementation of face recognition algorithm for biometrics based time attendance system. Retrieved from https://ieeexplore.ieee.org/abstract/document/7013165

Herrera, M. (2013). How to Perfect Your Risk Mitigation Strategies. Retrieved 26 August 2019, from https://www.mha-it.com/2013/05/17/four-types-of-risk-mitigation/

Heath, N. (2019). Raspberry Pi: A cheat sheet. Retrieved 26 August 2019, from https://www.techrepublic.com/article/raspberry-pi-the-smart-persons-guide/

Open Source Computer Vision. (2019). Retrieved 29 August 2019, from https://docs.opencv.org/3.4.3/d1/dfb/intro.html

Rosebrock, A. (2018). Face recognition with OpenCV, Python, and deep learning. Retrieved 29 August 2019, from https://www.pyimagesearch.com/2018/06/18/face-recognition-with-opencv-python-and-deep-learning/

Himanshi, E. (2015). Face Recognition Using Principal Component Analysis. Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7154774

Hartnett, M. (2015). Face Recognition: From Traditional to Deep Learning Methods. Retrieved from https://arxiv.org/pdf/1811.00116.pdf

Appendices

Appendix A – Research Project Plan

Task 2: Research Project Plan: Implementing Facial Recognition Technology Using a Microprocessor to Register Student Attendance

James Early
28040066

INTRODUCTION AND JUSTIFICATION

The goal for many schools and universities across the country is being able to reduce the absenteeism of its students. (Epstein, 2010) The issue is mainly related to the method in which their attendance is being captured. The most traditional approach taken is using the basic technique of pen and paper however in some more technologically advanced institutes, ID card scanning is being utilised. Marking a student's attendance is incredibly important as it provides staff an overview into the student's overall presence.

By combining existing technologies such as a microprocessor and facial recognition would allow for the construction of a unique system. Implementing a new system that requires minimal management and is automated will eradicate any possibility of producing inaccurate attendance data. The current process being used by Sheffield Hallam includes students manually recording their presence via a signature which in itself poses many potential issues such as missing out a student or even forgery. This only emphasises the fact that a new system is urgently needed.

RESEARCH QUESTION, AIMS & OBJECTIVES

Research question

How to quantify a student's attendance to produce accurate registration results?

Objectives

To envisage the answer to this question, the aims of the project must be met. The underlying aim is to create a process that is capable of capturing student's attendance in class. It needs to be reliable enough to supersede the current system and possess the ability to produce a robust set of results that can be used to monitor student's presence.

- Build a raspberry pi prototype with encompassing LCD screen and USB camera attachment
  S – Compact system that allows for easy demonstration
  M – LCD screen isn't essential but will provide a tidier deliverable
  A – The resources are cost-effective including the software development IDE is free.
  R – Essential to the projects overall success as it provides the computing power.
  T – Initial objective to accomplish as this will set the foundation of the project.

- Implement OpenCV and Deep Learning to enable facial recognition functionality
  S – Install the required pre-requisites to permit the use of facial recognition
  M – Both are required to output the results needed
  A – Precise modules and packages need to be installed in correct order to avoid corruption
  R – Essential to the projects overall success as it provides the facial detection.
  T – Required to be completed before GUI creation.

- Develop a Python GUI that allows for ease of use when operating system
  S – Provides a simple GUI to execute CMD functions
  M – Enough to fully test the system
  A – Python knowledge needs to be applied to created adequate GUI
  R – Non-essential however extremely pragmatic.
  T – Cosmetic objective therefore can be completed towards the end.

- Output a set of feasible results that could be later used for analytical purposes
  S – Export results gathered using MySQL tools
  M – Simple relational database to display basic results of attendance
  A – Prototype local database will be implemented to avoid hosting costs
  R – Added practicality to the project
  T – Minor time to build therefore can build towards the end.

Deliverable

The deliverable will be a fully functional Raspberry-Pi facial recognition system that has the ability to accurately simulate and display the results of a student's university class attendance.

LITERATURE REVIEW

An imperative element for the success of the education system is for the regular attendance of its students. This is inclusive for most higher education institutions however it can still be applied to lower level schools. One of the main issues today is getting students to attend their classes. This is becoming increasingly more frequent with university's as being present is no longer compulsory. This is a complete contrast for younger students who attend nursery and high-school as its more of a requirement than an option. That being said, many university's still have a benchmark of attendance expected from all students in the form of a percentage. If this isn't reached, then disciplinary action could occur.

This alone is still not enough to force students to attend therefore a certain degree of stimulation for a student needs to be applied. Traditionally, students should naturally understand that maintaining a constant presence in the majority of their classes will constitute to their overall grade. (Krishnan, 2015) In some scenarios, schools have started to use attendance points to promote the appearance of students which can then be used to reward or penalise them depending on how many they accumulate over a period of time. This method seems useful but still possess the risk of how the attendance is recorded.

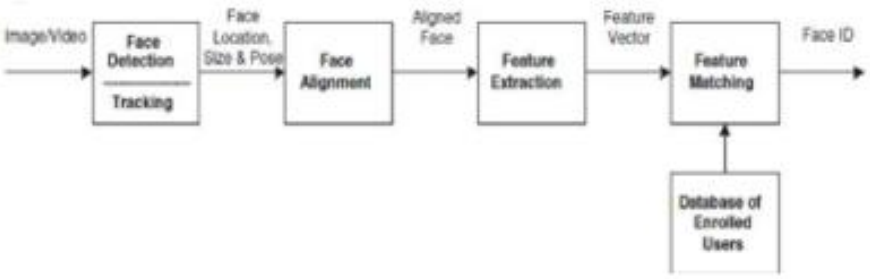
As discussed before, most modern systems manually record attendance either by using a paper-based register or scanning ID cards. This has proven to be a pragmatic approach as we still utilise this system today however it does contain its own set of issues. The most apparent drawback is that it adds an additional pressure onto the teacher to correctly mark the attendance of the present students which consequently also has the ability to disrupt the class. Other issues could include the possibility of missing out students entirely especially in a large classroom environment. (Krishnan, 2015)

Even though there are multiple issues relating to the current systems being used, this project focuses more into how we can record the data more accurately. By creating a robust method into capturing this type of data will eliminate any previous or subsequent issues identified with existing techniques. There are a magnitude of ways to capture the attendance of pupils whether it's using traditional methods or integrating technology. The most underused technique would include the incorporation of biometrics especially in an education environment. Biometrics such as facial recognition possess the merits of both accuracy and low-intrusiveness which would make it a perfect solution to replace current data collection methods. (Krishnan, 2015)

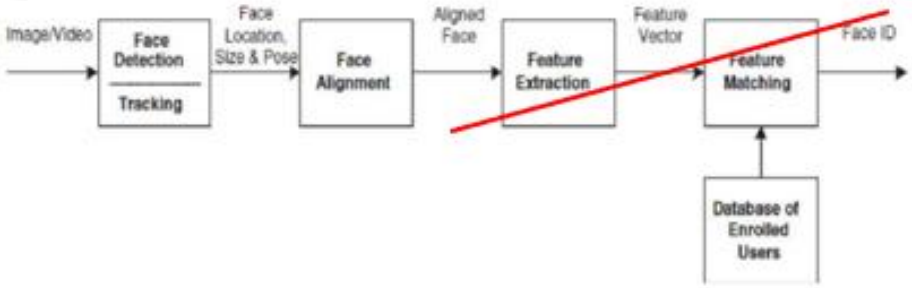
Using technology such as biometrics would prove to be much more effective however it does come with limitations. Implementing a system that utilises facial technology will warrant suitable hardware that is powerful enough to capture faces real time. In conjunction is also choosing an algorithm that is simple enough to work along side with the CPU without overloading it. Proposed facial recognition systems usually include multiple algorithms and adequate computing hardware capable of producing accurate readings from the students. "We have taken different techniques like colour based detection and Principle Component Analysis (PCA) for face detection and for feature extraction, PCA and Linear Discriminate Analysis (LDA)." (Jha, 2015)

We can see from Jha's implemented system that he uses the popular PCA and LDA algorithm that extracts the colours and features of the students face for increased facial detection results. This solution is the perfect answer for the research question however having the resources to execute and maintain this architecture will be challenging. Similarly, another proposed system uses the AdaBoost algorithm and SOC hardware framework to carry out their facial recognition (Fuzail, 2014) but again, implementing this setup using the resources available will not solve the complication identified.

The most substantial challenge of the is project is to build a system that able to solve the issue whilst utlitisng the resources avaiable. We have previously seen the requirements into how to get accurate results using high specification components however the difference is achieve the same results using lower level assets. In this instance, using a Raspberri-Pi microprosessor along with basic facial recognition algorithims such as the Haar Cascade and HOG detection provided by OpenCV and dlib libraries. They may not produce the highest results which would be achieved by using either Jha's or Fuzails implementations however it should still solve the ongoing issue that current student attendance systems present.



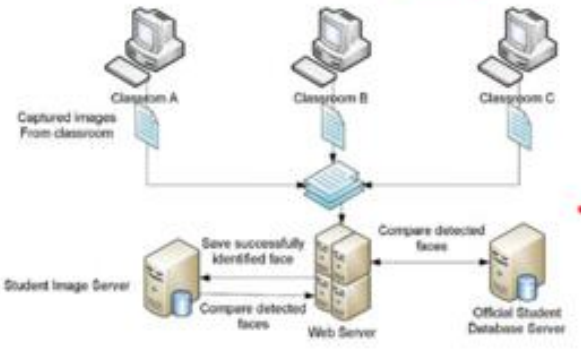
Flow of facial recognition algorithm (Jha, 2015)

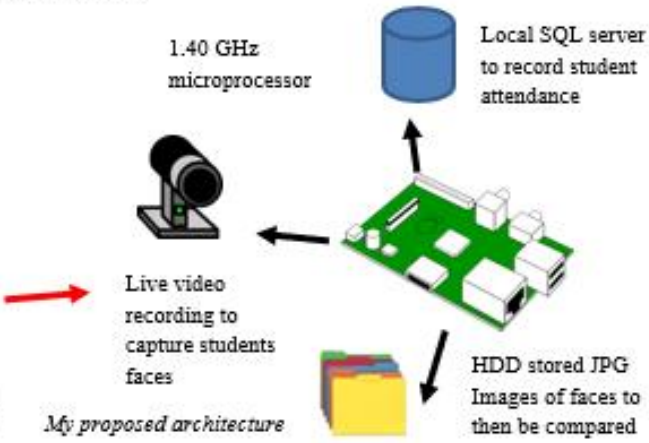


Comparison of Jha's algorithm to mine

Above is an illustration into the comparison between the algorithm that will be implemented to Jha's. Due to the processing power and GPU RAM capabilities between a microprocessor and Jha's architecture, feature detection will need to be disabled.

Fuzail's hardware architecture (Fuzail, 2014)





My proposed architecture

To give an idea into the architecture comparison, above you can see how the resources available differ from Fuzail's. Whereas the proposed solution contains all components being hosted on one

device, Fuzail's system contains multiple machines that each share the processing load. For this project to be a success, there needs to be the correct balance between choosing the correct algorithms that work effectively with the chosen hardware all while producing accurate results.

RESEARCH DESIGN

Starting a project of any kind will require copious amounts of planning before it commences. Research projects are no exception to this but their approach can differ from more traditional ventures. Typically, this type of project is conducted in a method that is set out to answer a specific question. Before the answer is identified, a series of stages will need to be considered in order to achieve the projects aims and objectives. One of the most substantial areas to cover includes the research design which should entail and demonstrate various elements of that particular project. It can be interpreted as a generic plan into how the research question will be answered.

In order to reach a successful outcome, implementing the research design is essential into achieving the desired outcome. When creating this proposal, its important to continuously take every step systematically to ensure there is a meaning behind the reason for conducting this type of research. The design forged will essentially influence the approach taken into answering the research question stated.

Research Philosophy

It's important to adopt the most suitable philosophy for the project as it provides a sense of realism. Researches who work on projects will naturally have their own views and beliefs into how the work should be conducted. This is usually shaped from previous experience and knowledge which will constitute how the research question is interpreted. The most adequate philosophy that can be delivered is the one of a pragmatist. This is because the data being collected will effect the actions that lead into answering the research question. This will only be the case if the collected data is credible and accurate.

Also, with this project there is no singular viewpoint that can be large enough to portray the bigger picture. As identified before there are alternative student attendance systems that exist therefore in another reality approaching the project differently could potentially provide more ideas into improving the proposed system. This state of mind benefits the project over other philosophies such as the realism. This is mainly because this mentality largely revolves around large datasets and includes qualitive data which the project doesn't contain.

Research Approach

For most research projects, there will be an incorporation into the use of theory. Some may approach the concept differently to others. Many projects begin with theory meaning they are able to conduct research effectively that is tailored around it. This is known as the deductive approach. They can also use this method to create a research strategy to test the theory. This project will explore the concept of both facial recognition and attendance systems therefore no new content is

being generated which the inductive approach would suggest. Data is simply being collected to modify existing theory which can then be tested. That being said, the abductive approach will be most suitable for the project.

Methodology

Choosing the correct methodology for your project is vital for the best outcome. The are a magnitude of approaches but the challenge is adopting the best one for your type of project. For this project, quantitative data will predominantly be collected from the subjects willing to participate. This will be in the form of JPG/PNG images of their faces for facial recognition testing. The data accumulated can then be used to determine the effectiveness of the system and provide an insight into the accuracy of the algorithm used. The reliability of the measurements are crucial as this could be then used as future work into improving the chosen algorithm or find alternative ways into producing better results. As an outcome of using the quantitative design, utilising the mono-method will be a suitable technique for the collection and analysis of data.

Research Strategy

As the project isn't part of a existing case study, it's imperative that the research strategy implemented caters for what it's trying to achieve. We already have identified that it isn't building a new solution but rather adapting current architectures to solve a specific problem. Therefore, adopting the feasibility study is ideal as its oriented around the idea of taking an existing system and discovering if its possible to make it work with the resources available. This strategy allows for the creation of a prototype which should be built around efficiency and effectiveness using the quantitative data gathered.

Data Collection and Analysis

Testing images of various quality will determine how accurately the system works. For example, capturing an image of a subject in both a dark and light setting to determine the precision of the facial detection system. Similarly, testing using subjects faces taken from different distances. The best approach would be to observe how the system performs under these circumstances and record the quality of the results. This data can then be analysed to see in exactly what scenarios the detection works best. Once the optimum setting for the data capture has been identified, this then will allow for the simulation of a live classroom environment. Other quantitative data can be recorded in the process such as how many faces can be recognised at once, the maximum fps the camera can run at, the duration a face can be recognised for etc.

ETHICS, RISKS AND ISSUES

The potential risks in this project are minimal as there are no dependencies on third party's or companies to acquire the data. All data being used is produced by the project from trusted participants. However, due to this nature, there is still a risk of communication. There is an expected level of constant participation from my subjects therefore If there is an issue then this could cause a problem to the projects progress. Other potential risks include technical as a substantial portion is working with both hardware and software which could malfunction at any

time. Mitigation strategies would be to accept the risks as new subjects for testing could be found or if it's a technical risk, buy new equipment as the essential components are cost effective.

The project isn't pushing any ethical or social boundaries but may have the potential to pose legal threats. Dealing with images that can be used to identify students warrants the possibility into breaching data protection acts. This means keeping their information secure at all times is paramount which can be done by using the appropriate encryption methods.

Resources

- Raspberri-Pi 3 Model B
- 64GB Micro SD card (Image, DB and OS storage)
- Microsoft HD 1080p Webcam
- Wireless Keyboard
- Wired Mouse
- 7inch LCD HDMI screen
- Rasbarian Linux OS system
- Python (GUI creation)
- OpenCV
- Dlib/face_recognition libraires
- Internet tutorials, articles and journals

Appendix B – Ethics Checklist Form

Sheffield Hallam University

RESEARCH ETHICS CHECKLIST FOR STUDENTS (SHUREC7)

This form is designed to help students and their supervisors to complete an ethical scrutiny of proposed research. The SHU Research Ethics Policy should be consulted before completing the form.

Answering the questions below will help you decide whether your proposed research requires ethical review by a Designated Research Ethics Working Group.

The final responsibility for ensuring that ethical research practices are followed rests with the supervisor for student research.

Note that students and staff are responsible for making suitable arrangements for keeping data secure and, if relevant, for keeping the identity of participants anonymous. They are also responsible for following SHU guidelines about data encryption and research data management.

The form also enables the University and Faculty to keep a record confirming that research conducted has been subjected to ethical scrutiny.

For student projects, the form may be completed by the student and the supervisor and/or module leader (as applicable). In all cases, it should be counter-signed by the supervisor and/or module leader, and kept as a record showing that ethical scrutiny has occurred. Students should retain a copy for inclusion in their research projects, and staff should keep a copy in the student file.

Please note if it may be necessary to conduct a health and safety risk assessment for the proposed research. Further information can be obtained from the Faculty Safety Co-ordinator.

General Details

Name of student	James Early
SHU email address	b8040066@my.shu.ac.uk
Course or qualification (student)	IT Management
Name of supervisor	Dr Elizabeth Uruchurtu
email address	E.Uruchurtu@shu.ac.uk
Title of proposed research	Implementing Facial Recognition Technology Using a Microprocessor to Register Student Attendance
Proposed start date	08/05/2019
Proposed end date	09/09/2019
Brief outline of research to include, rationale & aims (250-500 words).	Many University establishments are experiencing ongoing issues relating to the attendance of its students. This is mainly due to the method in which their attendance is captured. The most traditional approach taken is using the basic technique of pen and paper however in some more technologically advanced institutes, ID card scanning is being utilised. Marking a

	student's attendance is incredibly important as it provides staff an overview into the student's overall presence. I propose to combine existing technologies such as a microprocessor and facial recognition to build a unique prototype. Implementing a new system that requires minimal management and is automated will eradicate any possibility of producing inaccurate attendance data. The current process being used by Sheffield Hallam includes students manually recording their presence via a signature which in itself poses many potential issues such as missing out a student or even forgery. This only emphasises the fact that a new system is urgently needed. The underlying aim is to create a process that is capable of capturing student's attendance in class. It needs to be reliable enough to supersede the current system and possess the ability to produce a robust set of results that can be used to monitor student's presence. It's important to realise that this project is essentially building a pragmatic method into collecting a rigorous set of results through the medium of technology rather than investigating the wider issue of generic student registration systems.
Where data is collected from individuals, outline the nature of data, details of anonymisation, storage and disposal procedures if required (250-500 words).	The participants involvement will include an extraction of a singular image during the testing stage of the project. The image will portray a clear view of the face with no obstructions to ensure optimum recognition. The willing also need to show their faces during a live video feed directly from the camera to test the full accuracy of the prototype. The images can either be captured using the USB camera attached to the Raspberri-Pi or can be retrieved from external sources such as social media platforms e.g. their Facebook profile picture. All participants will be non-vulnerable whilst being made fully aware of the projects objectives and what's expected from them. To achieve accurate results, around 5 subjects will be approached to take part. These will be a mixture of personal friends and university colleagues. Prior to any involvement, all participants will be required to fill and sign a consent form which will demonstrate their understanding of the data being handled and their own responsibilities towards the project. Due to the nature of the data being collected, I will take extensive measures to ensure the anonymity of those taking part. The report itself will not contain any names of my subjects nor will it display the screenshots of the faces provided. In the scenario where I am required to use their data to illustrate my findings, alias names will be created, and images will be blurred. Image collection will be done privately using a secure pre-

	booked room. The images themselves will be stored inside a secure folder directory from within the memory card of the device. The device can only be accessed via a username and password providing an extra layer of protection. In addition, encryption of the folders containing any sensitive data will be executed using EcryptFS. The disposal of the data will include removing the files securely from the memory card. To ensure the eradication of all information created, the CMD line will be used to erase any potential hidden files and subfolders. It's imperative to understand that the outcome of this project will solely consist of a prototype system therefore It cannot be emphasized enough that it will not be implemented in a live classroom environment.

1. Health Related Research Involving the NHS or Social Care / Community Care or the Criminal Justice Service or with research participants unable to provide informed consent

Question	Yes/No
1. Does the research involve? • Patients recruited because of their past or present use of the NHS or Social Care • Relatives/carers of patients recruited because of their past or present use of the NHS or Social Care • Access to data, organs or other bodily material of past or present NHS patients • Foetal material and IVF involving NHS patients • The recently dead in NHS premises • Prisoners or others within the criminal justice system recruited for health-related research* • Police, court officials, prisoners or others within the criminal justice system* • Participants who are unable to provide informed consent due to their incapacity even if the project is not health related	No
2. Is this a research project as opposed to service evaluation or audit? *For NHS definitions please see the following website* http://www.hra.nhs.uk/documents/2013/09/defining-research.pdf	No

If you have answered **YES** to questions **1 & 2** then you **must** seek the appropriate external approvals from the NHS, Social Care or the National Offender Management Service (NOMS) under their independent Research Governance schemes. Further information is provided below.

NHS https://www.myresearchproject.org.uk/Signin.aspx

* All prison and probation projects also need HM Prison and Probation Service (HMPPS) approval. Further guidance at: https://www.myresearchproject.org.uk/help/hlphmpps.aspx

NB FRECs provide Independent Scientific Review for NHS or SC research and initial scrutiny for ethics applications as required for university sponsorship of the research. Applicants can use the NHS proforma and submit this initially to their FREC.

2. Research with Human Participants

Question	Yes/No
Does the research involve human participants? This includes surveys, questionnaires, observing behaviour etc.	Yes
Question	**Yes/No**
1. *Note If YES, then please answer questions 2 to 10* *If NO, please go to Section 3*	
2. Will any of the participants be vulnerable? *Note: Vulnerable' people include children and young people, people with learning disabilities, people who may be limited by age or sickness, etc. See definition on website*	No
3. Are drugs, placebos or other substances (e.g. food substances, vitamins) to be administered to the study participants or will the study involve invasive, intrusive or potentially harmful procedures of any kind?	No
4. Will tissue samples (including blood) be obtained from participants?	No
5. Is pain or more than mild discomfort likely to result from the study?	No
6. Will the study involve prolonged or repetitive testing?	No
7. Is there any reasonable and foreseeable risk of physical or emotional harm to any of the participants? *Note: Harm may be caused by distressing or intrusive interview questions, uncomfortable procedures involving the participant, invasion of privacy, topics relating to highly personal information, topics relating to illegal activity, etc.*	No
8. Will anyone be taking part without giving their informed consent?	No
9. Is it covert research? *Note: 'Covert research' refers to research that is conducted without the knowledge of participants.*	No
10. Will the research output allow identification of any individual who has not given their express consent to be identified?	No

If you answered **YES only** to question **1,** the checklist should be saved and any course procedures for submission followed. If you have answered **YES** to any of the other questions you are **required** to submit a SHUREC8A (or 8B) to the FREC. If you answered **YES** to question 8 and participants cannot provide informed consent due to their incapacity you must obtain the appropriate approvals from the NHS research governance system. Your supervisor will advise.

3. Research in Organisations

Question	Yes/No
1. Will the research involve working with/within an organisation (e.g. school, business, charity, museum, government department, international agency, etc.)?	Yes
2. If you answered YES to question 1, do you have granted access to conduct the research? *If YES, students please show evidence to your supervisor. PI should retain safely.*	No

3. If you answered NO to question 2, is it because: A. you have not yet asked B. you have asked and not yet received an answer C. you have asked and been refused access. *Note: You will only be able to start the research when you have been granted access.*	A

4. Research with Products and Artefacts

Question	Yes/No
1. Will the research involve working with copyrighted documents, films, broadcasts, photographs, artworks, designs, products, programmes, databases, networks, processes, existing datasets or secure data?	No
2. If you answered YES to question 1, are the materials you intend to use in the public domain? *Notes: 'In the public domain' does not mean the same thing as 'publicly accessible'.* • *Information which is 'in the public domain' is no longer protected by copyright (i.e. copyright has either expired or been waived) and can be used without permission.* • *Information which is 'publicly accessible' (e.g. TV broadcasts, websites, artworks, newspapers) is available for anyone to consult/view. It is still protected by copyright even if there is no copyright notice. In UK law, copyright protection is automatic and does not require a copyright statement, although it is always good practice to provide one. It is necessary to check the terms and conditions of use to find out exactly how the material may be reused etc.* *If you answered YES to question 1, be aware that you may need to consider other ethics codes. For example, when conducting Internet research, consult the code of the Association of Internet Researchers; for educational research, consult the Code of Ethics of the British Educational Research Association.*	
3. If you answered NO to question 2, do you have explicit permission to use these materials as data? *If YES, please show evidence to your supervisor.*	
4. If you answered NO to question 3, is it because: A. you have not yet asked permission B. you have asked and not yet received and answer C. you have asked and been refused access. *Note You will only be able to start the research when you have been granted permission to use the specified material.*	

Adherence to SHU policy and procedures

Personal statement
I can confirm that: – I have read the Sheffield Hallam University Research Ethics Policy and Procedures – I agree to abide by its principles.
Student

Name: James Early	Date: 08/05/2019
Signature: 	
Supervisor or other person giving ethical sign-off	
I can confirm that completion of this form has not identified the need for ethical approval by the FREC or an NHS, Social Care or other external REC. The research will not commence until any approvals required under Sections 3 & 4 have been received.	
Name: Dr Elizabeth Uruchurtu	Date: 08/05/2019
Signature:	
Additional Signature if required:	
Name:	Date:
Signature:	

Please ensure the following are included with this form if applicable, tick box to indicate:

	Yes	No	N/A
Research proposal if prepared previously	☐	☐	✓
Any recruitment materials (e.g. posters, letters, etc.)	☐	✓	☐
Participant information sheet	☐	☐	✓
Participant consent form	☐	☐	✓
Details of measures to be used (e.g. questionnaires, etc.)	☐	☐	✓
Outline interview schedule / focus group schedule	☐	✓	☐
Debriefing materials	☐	✓	☐
Health and Safety Project Safety Plan for Procedures	☐	✓	☐

Appendix C – SHUREC8

Sheffield Hallam University

APPLICATION FOR RESEARCH ETHICS APPROVAL FOR STUDENTS (SHUREC8A)

SECTION A: Research Protocol

Important Note - If you have already written a research proposal (e.g. for a funder) that answers the methodology questions in this section please include a copy of the proposal and leave those questions blank. You **MUST** however complete **ALL** of Section B and C (risk assessment).

1. **Student Name: James Early**

   Faculty: Science Technology and Arts (STA)

   Email address: b8040066@my.shu.ac.uk

2. **Title of research: Implementing Facial Recognition Technology Using a Microprocessor to Register Student Attendance**

3. **Supervisor: Dr Elizabeth Uruchurtu**

   Email address: aceseu@exchange.shu.ac.uk

4. **Proposed duration of project**

   Start date: 01/06/2019 **End Date: 16/09/2019**

5. **Location of research if outside SHU: N/A**

6. **Background to the study and scientific rationale** (500- 750 words approx.)

Many University establishments are experiencing ongoing issues relating to the attendance of its students. This is mainly due to the method in which their attendance is captured. The most traditional approach taken is using the basic technique of pen and paper however in some more technologically advanced institutes, ID card scanning is being utilised. Marking a student's attendance is incredibly important as it provides staff an overview into the student's overall presence.

I propose to combine existing technologies such as a microprocessor and facial recognition to build a unique prototype. Implementing a new system that requires minimal management and is automated will eradicate any possibility of producing inaccurate attendance data. The current process being used by Sheffield Hallam includes students manually recording their presence via a signature which in itself poses many potential issues such as missing out a student or even forgery. This only emphasises the fact that a new system is urgently needed.

The underlying aim is to create a process that is capable of capturing student's attendance in class. It needs to be reliable enough to supersede the current system and possess the ability to produce a robust set of results that can be used to monitor student's presence.

It's important to realise that this project is essentially building a pragmatic method into collecting a rigorous set of results through the medium of technology rather than investigating the wider issue of generic student registration systems.

7. Main research questions

How to quantify a student's attendance to produce accurate registration results?

8. Summary of methods including proposed data analyses

For this project, quantitative data will predominantly be collected from the subjects willing to participate. This will be in the form of JPG/PNG images of their faces for facial recognition testing. The data accumulated can then be used to determine the effectiveness of the system and provide an insight into the accuracy of the algorithm used. The reliability of the measurements are crucial as this could be then used as future work into improving the chosen algorithm or find alternative ways into producing better results. As an outcome of using the quantitative design, utilising the mono-method will be a suitable technique for the collection and analysis of data.

Testing images of various quality will determine how accurately the system works. For example, capturing an image of a subject in both a dark and light setting to determine the precision of the facial detection system. Similarly, testing using subjects faces taken from different distances. The best approach would be to observe how the system performs under these circumstances and record the quality of the results. This data can then be analysed to see in exactly what scenarios the detection works best. Once the optimum setting for the data capture has been exposed, this then will allow for the simulation of a live classroom environment. This includes an exercise that sees the participants walking into a room to simulate their attendance for class. The system will then run a live video feed that automatically executes the facial recognition algorithm to identify the students faces as they enter. It's important to understand that video or sound will not be recorded meaning this exercise will be entirely shot live in which observations will be made.

Other quantitative data can be noted in the process such as how many faces can be recognised at once, the maximum fps the camera can run at, the duration a face can be recognised for etc.

SECTION B: Ethics Proforma

1. **Describe the arrangements for selecting/sampling and briefing potential participants.** This should include copies of any advertisements for volunteers, letters to individuals/organisations inviting participation and participant information sheets. The sample sizes with power calculations if appropriate should be included.

The process to recruit potential participants will be accomplished by contacting them through either electronic messaging or approaching them in person. They will be chosen based on their similar academic interests and expertise therefore a mixture of close acquaintances and university colleagues would be ideal. They will be briefed using the information sheet attached before any data collection begins.

2. **What is the potential for participants to benefit from participation in the research?**

Knowing they have provided extensive help to improve the accuracy of the prototype system and validity of the results.

3. **Describe any possible negative consequences of participation in the research along with the ways in which these consequences will be limited.**

There is the possibility that the subject's sensitive data, such as facial images, will not be managed with due diligence. However, there will be strict security measures applied including storage encryption to prevent any data leakages. Participants maybe also be concerned that their data is being used for alternative purposes. This could be limited by providing them with regular updates of the project's progression in addition to proof that their data is being used to aid in the study.

4. **Describe the arrangements for obtaining participants' consent.** This should include copies of the information that they will receive & written consent forms where appropriate. If children or young people are to be participants in the study details of the arrangements for obtaining consent from parents or those acting in *loco parentis* or as advocates should be provided.

All participants will be required to fill and sign the consent form attached which demonstrates their understanding of the data being handled and what is expected from them. The form will reflect that the participant has acknowledged and accepted the conditions of the information sheet.

5. **Describe how participants will be made aware of their right to withdraw from the research.** This should also include information about participants' right to withhold information and a reasonable time span for withdrawal should be specified.

Information describing their rights to withdraw will be made abundantly clear both on the information sheet and consent form. I will also verbally reiterate their right to leave the experiment at any point during the project.

6. **If your project requires that you work with vulnerable participants describe how you will implement safeguarding procedures during data collection.**

Non-vulnerable participants will only be selected

7. **If Disclosure and Barring Service (DBS) checks are required, please supply details**

N/A

8. **Describe the arrangements for debriefing the participants.** This should include copies of the information that participants will receive where appropriate.

Subjects will be verbally informed of the outcome of the experiment and will be reassured that their data will be deleted upon project completion.

9. **Describe the arrangements for ensuring participant confidentiality.** This should include details of:
   - how data will be stored to ensure compliance with data protection legislation
   - how results will be presented
   - exceptional circumstances where confidentiality may not be preserved
   - how and when confidential data will be disposed of

The participants anonymity is paramount therefore security measures have been applied to maintain complete confidentiality during the longevity of the study.

Data storage – Facial images and DB records will be stored inside an encrypted folder. The backup hard drive will also be encrypted using the same algorithm.

Results – Results will be presented in the report however the participants identity will be protected as alias' will be used and images will be blurred.

Exceptional Circumstances – If the data was leaked or the device was misplaced, then potential hackers would have to not only break through the initial username and password authentication, they'd have to also crack the second layer of file encryption. If they did manage to access images of the participants, visual identity would only be compromised as an alias naming scheme will be implemented e.g. [photo 1] – "Participant A" etc.

Disposal of Data - The disposal of the data will include removing the files securely from the memory card and backup drive. To ensure the eradication of all information created, the CMD line will be used to erase any potential hidden files and subfolders.

10. **Are there any conflicts of interest in you undertaking this research?** (E.g. are you undertaking research on work colleagues or in an organisation where you are a consultant?) Please supply details of how this will be addressed.

No

11. **What are the expected outcomes, impacts and benefits of the research?**

A successful system in which to capture student's attendance using facial recognition

12. **Please give details of any plans for dissemination of the results of the research.**

Since the system is a prototype, there are no plans for disseminating the results.

SECTION C

HEALTH AND SAFETY RISK ASSESSMENT FOR THE RESEARCHER

1. Will the proposed data collection take place on campus?

✓ Yes (Please answer questions 4, 6 and 7)
☐ No (Please complete all questions)

2. Where will the data collection take place?
(Tick as many as apply if data collection will take place in multiple venues)

	Location	Please specify
☐	Researcher's Residence	
☐	Participant's Residence	
☐	Education Establishment	
☐	Other e.g. business/voluntary organisation, public venue	
☐	Outside UK	

3. How will you travel to and from the data collection venue?

☐ On foot ☐ By car ☐ Public Transport
☐ Other (Please specify)

Please outline how you will ensure your personal safety when travelling to and from the data collection venue

4. How will you ensure your own personal safety whilst at the research venue?

The study will take place within the university. This will provide a safe environment for the participants involved in the project.

5. If you are carrying out research off-campus, you must ensure that each time you go out to collect data you ensure that someone you trust knows where you are going (without breaching the confidentiality of your participants), how you are getting there (preferably including your travel route), when you expect to get back, and what to do should you not return at the specified time. Please outline here the procedure you propose using to do this.

6. Are there any potential risks to your health and wellbeing associated with either (a) the venue where the research will take place and/or (b) the research topic itself?

✓ None that I am aware of
☐ Yes (Please outline below including steps taken to minimise risk)

7. **Does this research project require a health and safety risk analysis for the procedures to be used?**

☐ Yes
✓ No

(If **YES** the completed Health and Safety Project Safety Plan for Procedures should be attached)

Adherence to SHU policy and procedures

Personal statement	
I confirm that: • *this research will conform to the principles outlined in the Sheffield Hallam University Research Ethics policy* • *this application is accurate to the best of my knowledge*	
Student	
Signature	
Date	*24/07/2019*
Supervisor	
Signature	
Date	
Other signature *(*if applicable*)*	
Signature	
Date	

Please ensure the following are included with this form if applicable, tick box to indicate:

	Yes	No	N/A
Research proposal if prepared previously	✓	☐	☐
Any recruitment materials (e.g. posters, letters, etc.)	✓	☐	☐
Participant information sheet	✓	☐	☐
Participant consent form	✓	☐	☐
Details of measures to be used (e.g. questionnaires, etc.)	✓	☐	☐
Outline interview schedule / focus group schedule	✓	☐	☐
Debriefing materials	✓	☐	☐
Health and Safety Project Safety Plan for Procedures	✓	☐	☐

Appendix D – Data Management Plan

Data Management Plan

Template for doctoral research projects

If your project funder requires a specific form of plan, please ensure that your plan meets the requirements. Please see the Library Research Support portal page and this DCC page for examples of funder specific plans.

1. What data will you collect or create?

Data collection will include a singular clear JPG image of the participant's face. If the participant is unavailable but still would like to take part, I will seek consent for the participant to provide a digital image of their face for testing purposes.

The main purpose of the testing is to determine the accuracy of the prototype in recognising the images obtained. This data will also be stored inside a local MySQL database.

2. How will your data be documented and described?

Both the facial images provided and screenshots taking during the testing stage will be used to illustrate my research and approach in my report. However, to preserve the anonymity of the participants, I will blur the images and use alias names to represent them.

3. How will your data be structured, stored, and backed up?

The structure consists of a local file directory which will be stored on a microSD card. The data itself will be encrypted using the EncryptFS algorithm for added security. All data including the file directory and database records will be backed up after each iteration of development on an external hard drive which will also be encrypted using the same algorithm.

RDM DMP Doctoral template v4 (30 January 2019)
Adapted from University of Cambridge (http://www.lib.cam.ac.uk/preservation/datatrain/documents.html) and University of Edinburgh (http://datalib.edina.ac.uk/mantra/Data_management_plan_template_MANTRA.docx)

4. How will you manage any ethical issues?

All ethical issues will be dealt with prior to the projects commence. The participants will be well informed of the nature of both the overall objectives and data being handled. They will be provided with an information sheet describing the projects aims and a consent form explaining their responsibilities during the experiment. It will be made evidentially clear that any participant involved can drop out at any point during the project.

5. What are your plans for data sharing after submission of your thesis?

The data produced will be minimal therefore I do not feel there will be any pragmatic use for my research. However, as discussed before, ensuring the anonymity of the participants could allow for my report to be published for educational purposes.

6. What are your plans for the long-term preservation of data supporting your research?

Due to the nature of data being captured and produced, the eradication of any potential sensitive information will be carried out.

Appendix E – Ethics Committee Approval

! STA Research Ethics Committee (FREC) <aces-frec-mb@exchange.shu.ac.uk>
to me, !, Elizabeth

Good morning James,

I'm pleased to inform you that the reviewer has recommended approval of your revised ethics submission.

Kind regards,

Alison Bremner
C3RI Administrator & PA to Dr Kathy Doherty
Cultural, Communication & Computing Research Institute (C3RI)
Faculty of Science, Technology and Arts (STA)
Room 9104, Cantor Building, Sheffield Hallam University, 153 Arundel Street, Sheffield, S1 2NU
Tel: 0114 225 4308 - Direct Line or Extension 4308

Appendix F – Information Sheet

Sheffield Hallam University

Student Attendance System Using Facial Recognition

Participant Information Sheet

The University undertakes research as part of its function for the community under its legal status. Data protection allows us to use personal data for research with appropriate safeguards in place under the legal basis of public tasks that are in the public interest. A full statement of your rights can be found at https://www.shu.ac.uk/about-this-website/privacy-policy/privacy-notices/privacy-notice-for-research. However, all University research is reviewed to ensure that participants are treated appropriately and their rights respected. This study was approved by UREC with Converis number ERxxxxxxx. Further information at https://www.shu.ac.uk/research/ethics-integrity-and-practice

Purpose of Study

I would like to invite you to take part in a study that aids into creating a protype capable of utilizing facial recognition technologies in an educational environment. The information you provide plays an integral role into reaching the overall aims and objectives of the endeavor. The research is primarily used to support the progression of the project whilst analyzing the data for system improvements.

Why have I been chosen?

You have been chosen based on our similar academic interests and expertise. As we share the same university course, I feel your attention to this study proves you are an adequate candidate to approach and will be an asset to the project.

Do I have to take part?

The decision to take part in the project is entirely down to you. This information sheet only serves the purpose of providing you details into what to expect. As such, you have the option to withdraw from the study at any given moment without providing a reason. This sheet is yours to keep and can be referred to at any time.

What will I be required to do?

You will be required to take part in 2 different exercises. The first involves an image being taken of your face which will be arranged in due course. The final stage is a live video session which includes participants simulating a classroom environment by walking past the camera as they enter the room to test the accuracy of the facial recognition system. This exercise can either take place individually or as a group depending on your personal preference. Either setting will produce adequate results.

Where will this take place?

Both exercises will take place inside a pre-booked room within the university. The date and time will be discussed with the participants individually and will be arranged based on their availability.

How often do I have to take part and for how long?

The initial stage of photo capturing will not last longer than 20 minutes. The live video session may last up to an hour as multiple recordings may be made. This is to ensure an adequate data set is produced for analysis purposes.

Are there any risks and/or disadvantages in taking part?

The most detrimental risk involved would include the leakage of your personal information. However, the security of your data is paramount therefore all security measures have been taken to ensure the safety of the material you provided. Mitigation plans have also been discussed to reduce the risk of any possible data related incidents. There are no disadvantages being part of this project as you are playing a vital role into the completion of the experiment.

Are there any benefits of taking part?

You are helping to gain a deeper understanding of facial recognition systems therefore the data you provide has the potential into creating new or improving similar technologies that could be seen in an educational environment in the near future.

When will I have the opportunity to discuss the participation?

There will be a chance for any questions about the work you have provided after each exercise. For example, a meeting could be arranged after both the image capture and live video sessions. Of course, questions are welcome during any stage of the project.

Will I be linked with this project in any way with what is recorded and reported?

Due to the photographic data being captured, your identity is of upmost importance to us therefore, we have taken extensive security measures to protect your sensitive data when storing it. Similarly, your anonymity will also be respected and maintained throughout the creation of the report as alias names will be generated when referring to any findings. E.g. "We found with Participant A". If images are used, then photo manipulation software will be utilised to thoroughly distort and blur any compromising information. Your confidentiality is taken seriously and will be enforced for the duration of the project.

Who will be responsible for the information collected when the study is over?

James Early is the lead researcher and owner of the project therefore he will be responsible for the data collected.

Who will have access to it?

Any sensitive information contributed by you will not be published therefore the project owner will only has access to it during the longevity of the study.

What will happen to the information when the study has finished?

All data provided by you will be eradicated from all systems and storage devices. This includes from the prototypes hard disk, any hard copies and from backup devices. Proof can be provided if required.

How will you use what you found out?

The results and observations made from the experiment will be illustrated in various chapters within the dissertation report. In no circumstances will the participants personal data be used in the literature.

How long will the study last?

The study and testing will likely last around a week. This includes the capturing of your data and testing of the prototype system.

How can I find out about the results?

The results will be made accessible upon request. They will be available when the study has concluded and been made adequate for distribution.

At any point during the research you feel the need to ask questions or seek extra clarification about any area of the study, then feel free to contact me on:

Email: b8040066@my.shu.ac.uk

This should be an enjoyable yet educational experience for both the participants and myself therefore it's vital to maintain a positive attitude and most importantly, have fun!

Thank you so much for taking part and helping me complete this project.

James Early – Lead Researcher and Project Owner

Any queries regarding the treatment of your personal data, use the contact details below:

You should contact the Data Protection Officer if:	**You should contact the Head of Research Ethics (Professor Ann Macaskill) if:**
• you have a query about how your data is used by the University • you would like to report a data security breach (e.g. if you think your personal data has been lost or disclosed inappropriately) • you would like to complain about how the University has used your personal data	• you have concerns with how the research was undertaken or how you were treated
DPO@shu.ac.uk	a.macaskill@shu.ac.uk
Postal address: Sheffield Hallam University, Howard Street, Sheffield S1 1WBT Telephone: 0114 225 5555	

Appendix G – Consent Forms

Sheffield Hallam University

Student Attendance System Using Facial Recognition

Participant Consent Form

Project Background:

This project involves the study into gaining a deeper understanding of how facial recognition systems work within a classroom environment. The objective is to create a prototype that is capable of recording student's attendance by scanning their face and marking them as being present in the room. The idea is to replace current inefficient solutions by implementing modern technology thus creating an improved and accurate system.

This sheet is for you to provide your consent into participating in the study. The details into what is required from you in addition to what can be expected from this project can be found in the **information sheet.**

Please answer the following questions by ticking one that applies to you:

		Yes	No
1.	I have read the information sheet and understand the details of the study.	☑	☐
2.	I have asked all the questions to recognise what is required of me and understand that I am allowed to ask further questions at any point during the project.	☑	☐
3.	I have acknowledged that I am allowed to withdraw from the study at any moment without providing a reason for my departure.	☑	☐
4.	I am happy and agree to attend the image capturing session to provide the researcher a facial image to be used for testing purposes.	☑	☐
5.	I am happy and agree to attend the live video session to aid in the research by simulating a classroom environment.	☑	☐
6.	I have read and understood the requirements stated on the information sheet and am happy to abide by them.	☑	☐
7.	I consent to the type of personal data collected from myself to be used for research purposes.	☑	☐
8.	I am aware that my identity will strictly be kept anonymous throughout the study and in the report.	☑	☐
9.	I understand that all traces of my data regarding my identity will be eradicated from all storage devices after the project's completion.	☑	☐
10.	I acknowledge that rigours security measures such as file encryption will be used to protect all my sensitive data.	☑	☐

Participant's Signature: _____ **Date:** 23/8/2019

Participant's Name (Printed): _____

Contact details: _____

_____

Researcher's Name (Printed): _____

Researcher's Signature: _____

Researcher's contact details:
(Name, address, contact number of investigator)

Please keep your copy of the consent form and the information sheet together.

Participant's Signature: _____ **Date:** 12 Aug 19

Participant's Name (Printed): _____

Contact details: _____

_____

Researcher's Name (Printed): _____

Researcher's Signature: _____

Researcher's contact details:
(Name, address, contact number of investigator)

Please keep your copy of the consent form and the information sheet together.

Participant's Signature: _____ Date: 6/9/19

Participant's Name (Printed): _____

Contact details: _____

_____

Researcher's Name (Printed): _____

Researcher's Signature: _____

Researcher's contact details:
(Name, address, contact number of investigator)

Please keep your copy of the consent form and the information sheet together.

Participant's Signature: _____ Date: 7/9/19

Participant's Name (Printed): _____

Contact details: _____

_____

Researcher's Name (Printed): _____

Researcher's Signature: _____

Researcher's contact details:
(Name, address, contact number of investigator)

Please keep your copy of the consent form and the information sheet together.

Participant's Signature: D. Crawley **Date:** 7/9/19

Participant's Name (Printed): _____

Contact details: _____

_____

Researcher's Name (Printed): _____

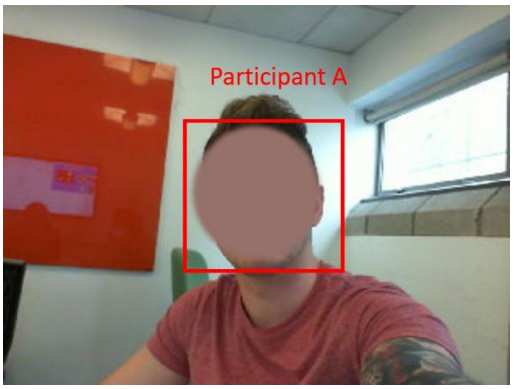
Researcher's Signature: _____

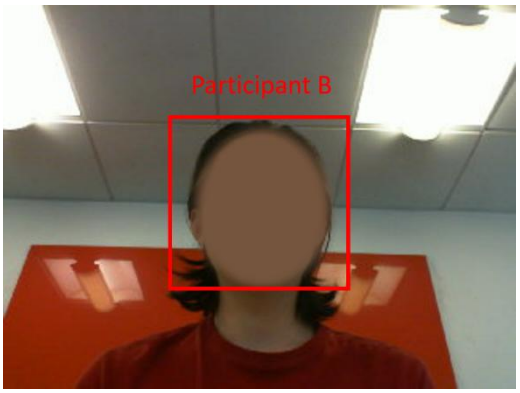
Researcher's contact details:
(Name, address, contact number of investigator)

Please keep your copy of the consent form and the information sheet together.

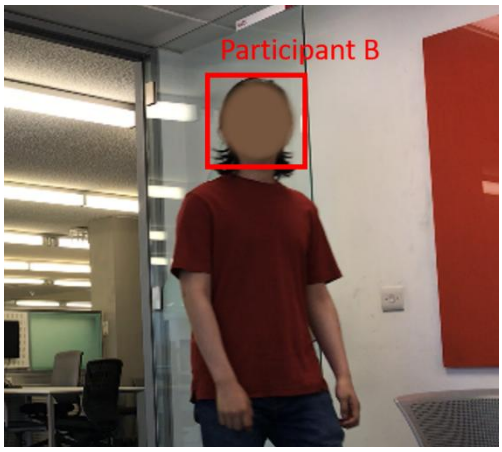
Appendix H – Results of Exercises

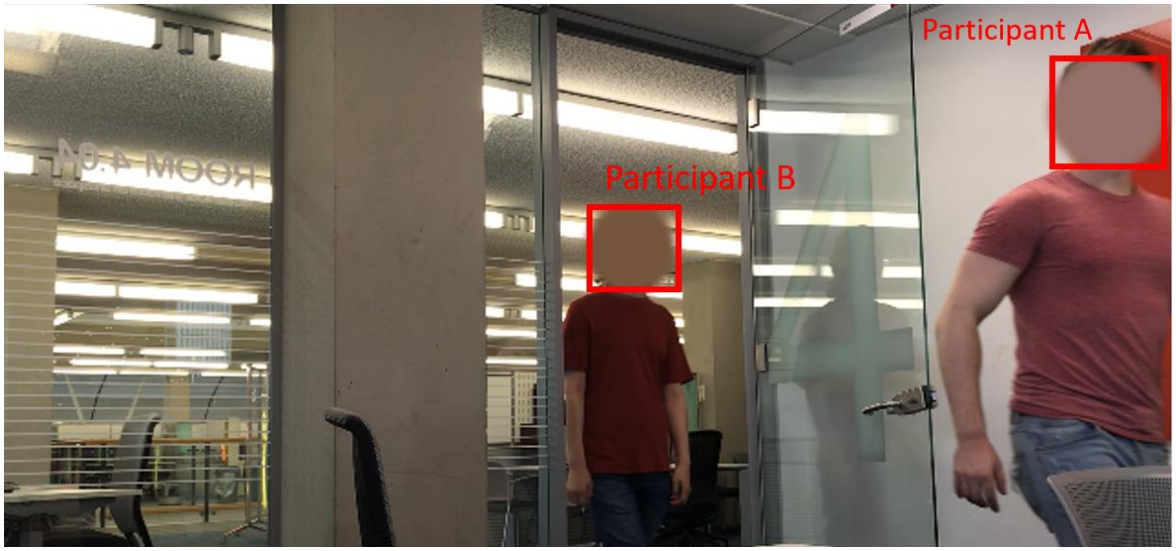
Exercise 1 – Still Image Comparison





Exercise 2 – Live Video Feed Detection Results





Appendix I – Python GUI Code

```
from tkinter import *
import os
import subprocess
from tkinter import messagebox
from tkinter.simpledialog import askstring
from tkinter.messagebox import showinfo

# creating tkinter window
root = Tk()
root.geometry('800x480')
root.title("Student Attendnace System")
root.configure(background='Blue')

def takePicture():
  myvar = askstring('Name', 'Name of Participant?')
  if not os.path.exists('dataset\%s'(myvar)):
    os.mkdir('dataset\%s'(myvar))
    os.system("fswebcam -r 1280x720 --no-banner dataset\%s"(myvar))
    messagebox.showinfo( "Image Capture", "New Participant Created")
  else:
    messagebox.showinfo( "Image Capture", "Participant Already Exsists")

  #entry = Entry(root)
  #entry.grid(column=6,row=2)
  #var1 = entry.get()

  #save = Button(root, command = testLiveFeed)
  #save.grid(column=6, row=2)

def encodeImages():
  os.system("virtualenvbackup/new/cv/bin/python encode_faces.py --dataset dataset --
encodings encodings.pickle --detection-method hog")
```

```
    messagebox.showinfo( "Encode Images", "New Images Encoded Successfully")


def testLiveFeed():

    os.system("virtualenvbackup/new/cv/bin/python recognize_face_live_video.py --cascade
haarcascade_frontalface_default.xml --encodings encodings.pickle")


def stillImage():

    os.system("virtualenvbackup/new/cv/bin/python recognize_faces_image.py --encodings
encodings.pickle --detection-method hog --image examples/example_01.jpg")


#def save():

   # os.system("/home/pi/.virtualenvs/cv/bin/python recognize_faces_image.py --encodings
encodings.pickle --detection-method hog --image examples/example_01.jpg")


heading = Label(root, text="Student Attendance System", fg='red')

heading.config(font=("Ariel", 20))

heading.grid(row=0, columnspan=20, pady=20)


imagetest = PhotoImage(file="GUIimages/2/camera.png")

btn = Button(root, image=imagetest, command = takePicture)

btn.grid(column=0, row=1,)


imagetest1 = PhotoImage(file="GUIimages/2/greenplus.png")

btn1 = Button(root, image=imagetest1, command = encodeImages)

btn1.grid(column=1, row=1)


imagetest2 = PhotoImage(file="GUIimages/2/live.png")

btn2 = Button(root, image=imagetest2, command = testLiveFeed)

btn2.grid(column=0, row=2)


imagetest3 = PhotoImage(file="GUIimages/2/picture.png")

btn3 = Button(root,image=imagetest3, command = stillImage)

btn3.grid(column=1, row=2)


imagetest4 = PhotoImage(file="GUIimages/2/picture.png")
```

```
btn4 = Button(root,image=imagetest3, command = stillImage)
btn4.grid(column=6, row=1, padx=100)

mainloop()
```

Appendix J – Facial Recognition Script (Still Image Comparison Exercise)

```
import face_recognition
import argparse
import pickle
import cv2
import MySQLdb

# construct the argument parser and parse the arguments
ap = argparse.ArgumentParser()
ap.add_argument("-e", "--encodings", required=True,
   help="path to serialized db of facial encodings")
ap.add_argument("-i", "--image", required=True,
   help="path to input image")
ap.add_argument("-d", "--detection-method", type=str, default="cnn",
   help="face detection model to use: either `hog` or `cnn`")
args = vars(ap.parse_args())

# load the known faces and embeddings
print("[INFO] loading encodings...")
data = pickle.loads(open(args["encodings"], "rb").read())

# load the input image and convert it from BGR to RGB
image = cv2.imread(args["image"])
rgb = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)

# detect the (x, y)-coordinates of the bounding boxes corresponding
# to each face in the input image, then compute the facial embeddings
# for each face
print("[INFO] recognizing faces...")
boxes = face_recognition.face_locations(rgb,
   model=args["detection_method"])
encodings = face_recognition.face_encodings(rgb, boxes)
```

```
# initialize the list of names for each face detected
names = []

# loop over the facial embeddings
for encoding in encodings:
    # attempt to match each face in the input image to our known
    # encodings
    matches = face_recognition.compare_faces(data["encodings"],
        encoding)

    name = "Unknown"

    # check to see if we have found a match
    if True in matches:
        # find the indexes of all matched faces then initialize a
        # dictionary to count the total number of times each face
        # was matched
        matchedIdxs = [i for (i, b) in enumerate(matches) if b]
        counts = {}

        # loop over the matched indexes and maintain a count for
        # each recognized face face
        for i in matchedIdxs:
            name = data["names"][i]
            counts[name] = counts.get(name, 0) + 1

        # determine the recognized face with the largest number of
        # votes (note: in the event of an unlikely tie Python will
        # select first entry in the dictionary)
        name = max(counts, key=counts.get)



    # update the list of names
```

```
    names.append(name)
    global c
    global db

    if __name__ == '__main__':
        try:
            db = MySQLdb.connect("localhost","root","root","studentattendancesystem")
            c= db.cursor()
        except:
            print ("No Connection to Server...")

        try:
          main()
        except KeyboardInterrupt:
          print ("bye bye...")
          pass

# loop over the recognized faces
for ((top, right, bottom, left), name) in zip(boxes, names):
    # draw the predicted face name on the image
    cv2.rectangle(image, (left, top), (right, bottom), (0, 255, 0), 2)
    y = top - 15 if top - 15 > 15 else top + 15
    cv2.putText(image, name, (left, y), cv2.FONT_HERSHEY_SIMPLEX,
        0.75, (0, 255, 0), 2)

# show the output image
cv2.imshow("Image", image)
cv2.waitKey(0)
```

Appendix K - Facial Recognition Script (Live Video Feed Exercise)

```
from imutils.video import VideoStream
from imutils.video import FPS
import face_recognition
import argparse
import imutils
import pickle
import time
import cv2
import MySQLdb

# construct the argument parser and parse the arguments
ap = argparse.ArgumentParser()
ap.add_argument("-c", "--cascade", required=True,
   help = "path to where the face cascade resides")
ap.add_argument("-e", "--encodings", required=True,
   help="path to serialized db of facial encodings")
args = vars(ap.parse_args())

# load the known faces and embeddings along with OpenCV's Haar
# cascade for face detection
print("[INFO] loading encodings + face detector...")
data = pickle.loads(open(args["encodings"], "rb").read())
detector = cv2.CascadeClassifier(args["cascade"])

# initialize the video stream and allow the camera sensor to warm up
print("[INFO] starting video stream...")
vs = VideoStream(src=0).start()
# vs = VideoStream(usePiCamera=True).start()
time.sleep(2.0)
# start the FPS counter
fps = FPS().start()
```

```
# loop over frames from the video file stream
while True:
    # grab the frame from the threaded video stream and resize it
    # to 500px (to speedup processing)
    frame = vs.read()
    frame = imutils.resize(frame, width=400)

    # convert the input frame from (1) BGR to grayscale (for face
    # detection) and (2) from BGR to RGB (for face recognition)
    gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
    rgb = cv2.cvtColor(frame, cv2.COLOR_BGR2RGB)

    # detect faces in the grayscale frame
    rects = detector.detectMultiScale(gray, scaleFactor=1.1,
        minNeighbors=5, minSize=(30, 30),
        flags=cv2.CASCADE_SCALE_IMAGE)

    # OpenCV returns bounding box coordinates in (x, y, w, h) order
    # but we need them in (top, right, bottom, left) order, so we
    # need to do a bit of reordering
    boxes = [(y, x + w, y + h, x) for (x, y, w, h) in rects]

    # compute the facial embeddings for each face bounding box
    encodings = face_recognition.face_encodings(rgb, boxes)
    names = []

    # loop over the facial embeddings
    for encoding in encodings:
        # attempt to match each face in the input image to our known
        # encodings
        matches = face_recognition.compare_faces(data["encodings"],
            encoding)
        name = "Unknown"
        # check to see if we have found a match
```

```
        if True in matches:
            # find the indexes of all matched faces then initialize a
            # dictionary to count the total number of times each face
            # was matched
            matchedIdxs = [i for (i, b) in enumerate(matches) if b]
            counts = {}

            # loop over the matched indexes and maintain a count for
            # each recognized face face
            for i in matchedIdxs:
                name = data["names"][i]
                counts[name] = counts.get(name, 0) + 1

            # determine the recognized face with the largest number
            # of votes (note: in the event of an unlikely tie Python
            # will select first entry in the dictionary)
            name = max(counts, key=counts.get)

        # update the list of names
        names.append(name)
        global c
        global db

        def insert_to_db():
            tme = time.strptime('my date', "%b %d %Y %H:%M")
            sql = "UPDATE ClassA SET Present = 'Yes, Time = (%s) WHERE FirstName =
(%s)"
            try:
                c.execute(sql, (time.strftime('%Y-%m-%d %H:%M:%S', tme)), [name])
                db.commit()
            except (MySQLdb.Error, MySQLdb.Warning) as e:
                print (e)
                db.rollback()
            #db.close()
```

```
    def main():
        insert_to_db()

    if __name__ == '__main__':
        try:
            db = MySQLdb.connect("localhost","root","root","studentattendancesystem")
            c= db.cursor()
        except:
            print ("No Connection to Server...")

        try:
          main()
        except KeyboardInterrupt:
          print ("bye bye...")
          pass
# loop over the recognized faces
for ((top, right, bottom, left), name) in zip(boxes, names):
    # draw the predicted face name on the image
    cv2.rectangle(frame, (left, top), (right, bottom),
        (0, 255, 0), 2)
    y = top - 15 if top - 15 > 15 else top + 15
    cv2.putText(frame, name, (left, y), cv2.FONT_HERSHEY_SIMPLEX,
        0.75, (0, 255, 0), 2)

# display the image to our screen
cv2.imshow("Frame", frame)
key = cv2.waitKey(1) & 0xFF
# if the `q` key was pressed, break from the loop
if key == ord("q"):
    break
# update the FPS counter
fps.update()
```

```
# stop the timer and display FPS information
fps.stop()
print("[INFO] elasped time: {:.2f}".format(fps.elapsed()))
print("[INFO] approx. FPS: {:.2f}".format(fps.fps()))

# do a bit of cleanup
cv2.destroyAllWindows()
vs.stop()
```

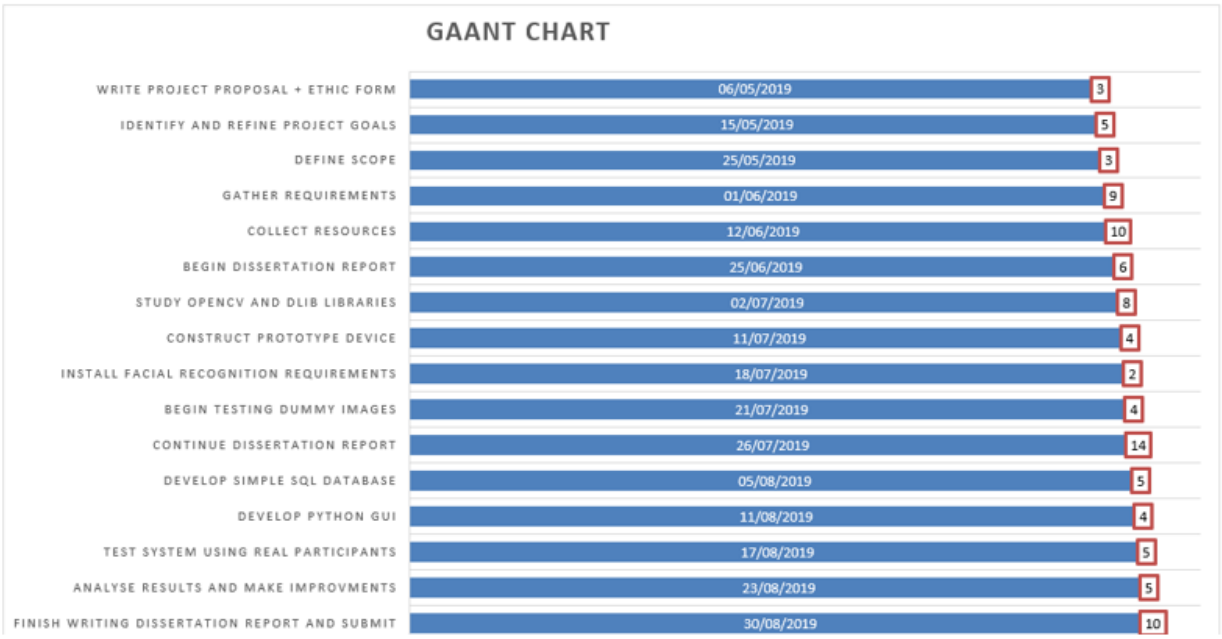
Appendix L – MySQL Database Creation Script (Exported)

```
-- phpMyAdmin SQL Dump
-- version 4.8.4
-- https://www.phpmyadmin.net/
--
-- Host: 127.0.0.1:3306
-- Generation Time: Sep 11, 2019 at 03:19 PM
-- Server version: 5.7.24
-- PHP Version: 7.2.14

SET SQL_MODE = "NO_AUTO_VALUE_ON_ZERO";
SET AUTOCOMMIT = 0;
START TRANSACTION;
SET time_zone = "+00:00";


/*!40101 SET @OLD_CHARACTER_SET_CLIENT=@@CHARACTER_SET_CLIENT */;
/*!40101 SET @OLD_CHARACTER_SET_RESULTS=@@CHARACTER_SET_RESULTS
*/;
/*!40101 SET @OLD_COLLATION_CONNECTION=@@COLLATION_CONNECTION
*/;
/*!40101 SET NAMES utf8mb4 */;

--
-- Database: `studentattendancesystem`
--

-- --------------------------------------------------------

--
-- Table structure for table `classa`
--

DROP TABLE IF EXISTS `classa`;
```

```
CREATE TABLE IF NOT EXISTS `classa` (
  `StudentID` int(10) NOT NULL AUTO_INCREMENT,
  `FirstName` varchar(50) DEFAULT NULL,
  `Surname` varchar(50) DEFAULT NULL,
  `Present` varchar(10) DEFAULT NULL,
  PRIMARY KEY (`StudentID`)
) ENGINE=MyISAM AUTO_INCREMENT=6 DEFAULT CHARSET=latin1;

--
-- Dumping data for table `classa`
--

INSERT INTO `classa` (`StudentID`, `FirstName`, `Surname`, `Present`) VALUES
(1, 'Participant A', NULL, NULL),
(2, 'Participant B', NULL, NULL),
(3, 'Participant C', NULL, NULL),
(4, 'Participant D', NULL, NULL),
(5, 'Participant E', NULL, NULL);
COMMIT;

/*!40101 SET CHARACTER_SET_CLIENT=@OLD_CHARACTER_SET_CLIENT */;
/*!40101 SET CHARACTER_SET_RESULTS=@OLD_CHARACTER_SET_RESULTS */;
/*!40101 SET COLLATION_CONNECTION=@OLD_COLLATION_CONNECTION */;
```

Appendix M – GANTT Chart

TIME PLAN



Task Name	Start	End	Duration (days)
Write Project Proposal + Ethic Form	06/05/2019	09/05/2019	3
Identify and Refine Project Goals	15/05/2019	20/05/2019	5
Define Scope	25/05/2019	28/05/2019	3
Gather Requirements	01/06/2019	10/06/2019	9
Collect Resources	12/06/2019	22/06/2019	10
Begin Dissertation Report	25/06/2019	01/07/2019	6
Study OpenCV and Dlib Libraries	02/07/2019	10/07/2019	8
Construct Prototype Device	11/07/2019	15/07/2019	4
Install Facial Recognition Requirements	18/07/2019	20/07/2019	2
Begin Testing Dummy Images	21/07/2019	25/07/2019	4
Continue Dissertation Report	26/07/2019	05/08/2019	14
Develop Simple SQL Database	05/08/2019	10/08/2019	5
Develop Python GUI	11/08/2019	15/08/2019	4
Test System Using Real Participants	17/08/2019	22/08/2019	5
Analyse Results and Make Improvments	23/08/2019	28/08/2019	5
Finish Writing Dissertation Report and Submit	30/08/2019	09/09/2019	10