SURVEILLANCE BASED CHILD KIDNAP DETECTION AND PREVENTION ASSISTANCE

2021-115

Final Report

K.A.O.V Kodikara

B.Sc. (Hons) Degree in Information Technology Specialized in Data Science

Department of Information Technology

Sri Lanka Institute of Information Technology Sri Lanka

October 2021

SURVEILLANCE BASED CHILD KIDNAP DETECTION AND PREVENTION ASSISTANCE

2021-115

Final Report

B.Sc. (Hons) Degree in Information Technology Specialized in Data Science

Department of Information Technology

Sri Lanka Institute of Information Technology Sri Lanka

October 2021

DECLARATION PAGE OF THE CANDIDATES & SUPERVISOR

"I declare that this is my own work and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. Also, I hereby grant to Sri Lanka Institute of Information Technology, the nonexclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Name	Student ID	Signature
K.A.O.V Kodikara	IT18110708	

The supervisor/s should certify the proposal report with the following declaration.

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

……………………………   …………………..

Signature of the supervisor: Date

ABSTRACT

The issue of children household accidents, kidnapping, and getting injured due to harmful child activities has raised significant concerns of public interest. The advancement in computer vision, particularly in image classification and object detection, could be applied to overcome the current flaws of the harmful activities detection devices that often failed to serve as a triggering system for guardians to protect their children from household accidents.

A real-time harmful child activity detection and prevention sssistance system that is mainly consisted of cameras as an input medium, a classifier to detect the presence of a child and a triggering system in audio and visual forms is a perfect utilized tool that aids in child protection. Monitoring systems on children’s behavior in closed spaces is possible to recognize their activity by the cameras in the environment, collect data, and then using learning algorithms. A deep learning algorithm to recognize and classify children's activity in smart homes is proposed with data collected from cameras in a smart environment, as the overall group research.

This paper is based on “Child Kidnap Detection and Prevention” to identify susceptible child kidnap by unauthorized persons. The intelligent surveillance system proposed for this is known as ‘AICare’. Purpose behind developing a proper kidnap detection methodology is to enhance and strengthen the existing child security systems. The key is to identify the main characteristics of a kidnapper in real time which follows face authentication, motion speed detection and object detection theories. Face detection is used to identify whether the outlined individual has covered his body specially face in a way of hiding the true identity, or else the person’s face is directly processed to face recognition for authorization. Motion speed detection is useful in calculating the average speed of movement of the targeted individual. Finally, the stranger is subjected to object detection in order to classify whether the person is holding a knife in hand. The captured outcomes are subjected to a decision tree to resolve the person as a kidnapper suspect.

Keywords: intelligent surveillance system, kidnap detection, video processing, real-time

ACKNOWLEDGEMENT

I would like to express our appreciation and full-hearted gratitude to our supervisor Mr. Prasanna Sumathipala and co-supervisor Ms. Sandamali Wijekoon who guide us to success this research and for sharing the experiences and expertise with the project matters. We also extend our gratitude to the Sri Lanka Institute of Information Technology (SLIIT) for giving an opportunity to carry out this research project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Description
SPP	Spatial Pyramid Pooling
PAN	Path Aggregation Network

LIST OF APPENDICES

1. INTRODUCTION

Kidnap detection has a large number of applications such as: security, area localizing, internet communication, and pose detection. The problem of kidnap detection is still unsolved and offers a great challenge to researchers because none of these researches address the main problem of abduction in human children. There are researches which depicts kidnap detection and prevention aid centralizing a robot's pose and in a limited area space [1]. In this situation kidnap detections for real-time abductions have come into need.

Facial authentication is a form of face recognition where faces captured in an image, or a video stream is compared against a known database of faces to verify identity. Face authentication is popularly used in the areas such as forensic investigations, protect law enforcement and smarter advertising. Here it depicts the significant landmark of face authentication in kidnap detection and child protection for recent years [2].

A kidnapper is a person of quick movements and is someone who is anxious on making the abduction quickly and smoothly as possible. Thereby, capturing these quick movements before the danger happens is something that is vital to be done in a kidnap detection system. The system should be capable to capture the speed of motion of the suspect in order to proceed with conclusions. To fulfil this requirement speed detection methodologies are used. Though speed detection work in vehicles is visibly abundant, motion speed detectors for humans are sparse.

An individual being hold of sharp objects in hand can lead him to be a kidnapper suspect or more. Object detection accurizes the goal of identifying a kidnapper suspect in real time as quickly and accurate as possible before any danger occurs. Object detection locate the position of the sharp object and identify whether it is being held by the person.

Children safety is the prime factor of concern to their parents. Despite the best safety measures, children, due to their lack of skills to protect themselves, may end up in a situation that endangers their life. In this paper, it is focused on the risk associated to household kidnapping when children are allowed to wander in their room alone. There have been previous incidents where children have been abducted from their foster home[3]. To improve child safety some parents, employ a nanny to look after the children inside the house. Nonetheless, human oversight absence may still lead to heartbreaking ending as in the previous cited story. This paper presents a real-time kidnap detection system to monitor the presence of unauthorized persons in the room with the child to enhance overall safety to the child from a possible kidnap.

1.1 Research Contributions

This paper provides the following contribution to the research community:

1. Provides a procedure to detect person holdings an object in hand.
2. Provides a mode of operation to detect speed of motion in humans.
3. Provides a methodology to detect suspiciously covered faces.
4. Provides a framework to detect child kidnap.
5. Provides baseline techniques to prevent kidnap before it occurs.
6. Provides methodology to detect kidnapper suspect using a sequence of detection results based on kidnapper characteristics.

1.2 Background and Literature Survey

This section presents the most related work to the problem addressed in this paper. A system to track children using a child module that transmits the tracking information to a database and a mobile device is proposed in[4]. The disadvantages of this system are deployment is expensive and the module may not be convenient for children.

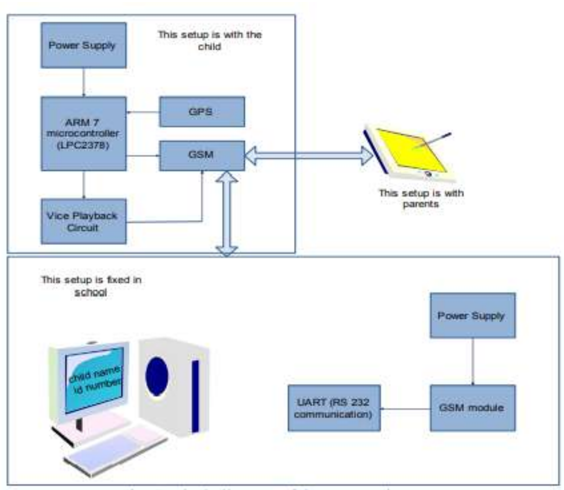


Figure 1: System design diagram of child tracking system[4]

[5] is about a tracking system that utilizes Android terminals that communicate among themselves using Bluetooth technology to form clusters. The clusters communicate that relevant information using WLAN. The major drawback of this system is that the deployment cost is very high.

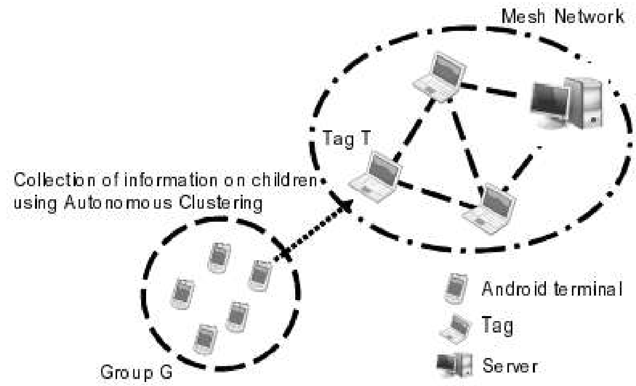


Figure 2: Children tracking system using Android mobile terminals[5]

There are commercial systems for tracking children such as Bluetooth-based tracking devices which are designed to be worn by children as a bracelet or a necklace[6]. In these tracking systems, these devices can be connected with a mobile application and if the child went outside a range specified by the parents, it can alert them. If the child was moved outside this range, the device will send an alert to the parent. In addition, the application sends the location of the child by using a geographical map. One disadvantage of this type of applications is that they notify only when the child has been moved out of the safety zone, in which the danger might have already happen.

Other products may rely on biometric features such as Fingerprint Recognition for children in which the children's scanned fingerprints are obtained using an acquisition device[7]. It uses image processing to enhance the information on the image and feature extraction to encode the found information. Then, the images are sent for comparison against a secure database of pre-registered users' patterns. Based on implementing this, it is possible to administer the entrance and leave times of a child from a particular place. The disadvantage of this approach is that it is not automatic and difficult for young children to use the acquisition device and correctly place their fingers on it. This may lead to inaccurate identifications and verifications.

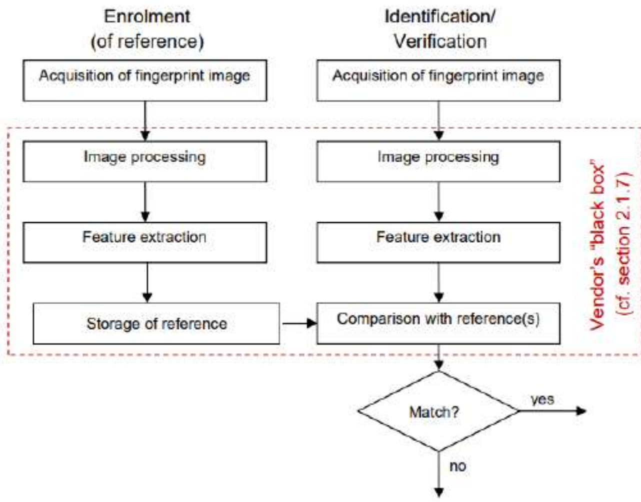


Figure 3: Fingerprint recognition for children[7]

Intelligent security camera concept is an area that is widely subjected to research and knowledge.[8] has developed an intelligent security camera system for kidnap detection. This study has proposed a system that can detect kidnapping cases using vehicles and automatically send an alert message to emergency services. This paper introduces a function to automatically detect criminal behaviors into the security system of the conventional human monitoring-based system. This function realizes efficient monitoring through many security cameras, and in addition to image recording, it effectively achieves crime prevention and immediate response. The main disadvantage of this system is that it only detects kidnappings that take place from vehicles.

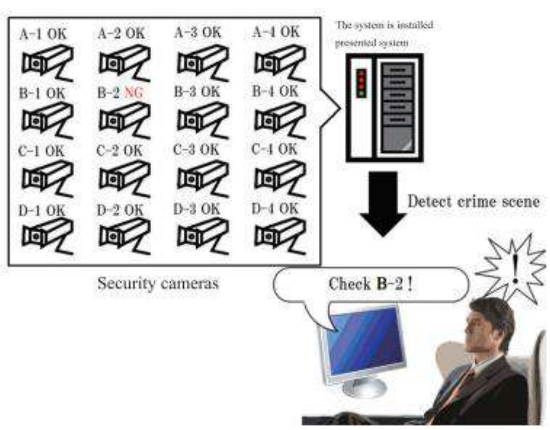


Figure 4: Overview of intelligent security camera system for kidnap detection[8]

[9] is a paper that explains a kidnapping detection scheme in which human pose estimation is used to classify accurately between kidnapping cases and normal ones. To estimate human poses from input video, human's 10 joint information is extracted by OpenPose library. The human pose estimation features which are computed from the location of detected human's joints are used as the features to distinguish kidnapping situations from the normal accompanying ones. The study that is proposed in this paper does not appear to be reliable for child abduction. As small children of the preschool age don't have to be forced taken for a kidnap to occur. These kids may follow the instructions of the kidnapper as small children have a tendency to listen to whatever an adult instructs. The methodology proposed by this paper is unable of capturing such situations.

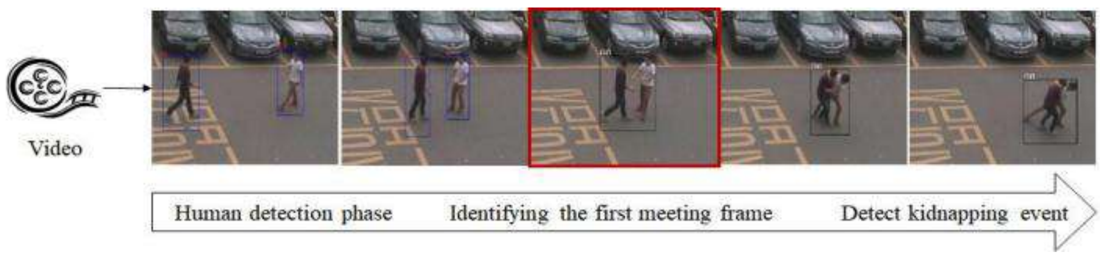


Figure 5: Kidnap detection process using human pose estimation[9]

Same as the above mentioned study[10] also proposes a study to develop a kidnapping event detection scheme for intelligent video surveillance by frame-based classification which is able to assort each frame into a kidnapping or normally accompanying situation. In this study, for generating training data from videos, a semi-automatic video annotation tool named INHA-VAT is used. Also a frame-based event classifier using Bayesian network model was developed to distinguish the frame of kidnapping situations from one of accompanying ones. Nonetheless child kidnapping situations by accompanying cannot be prevented using this method.[11] says that children are very open and trusting and therefore they are prone to kidnappers who disguise as trustworthy.

According to [1] kidnapping is a localization problem that occurs when an unexpected movement happens to the target due to interaction with its surroundings. It causes incorrect pose estimation of the target. This study proposes a method to solve this problem in the previous known-map situation in robots. The authors assume that the robot is kidnapped to the explored area during SLAM and is relocalized successfully. Kidnap detections built centralizing robots consider different features than of when detecting child kidnap. Therefore this methodology cannot be implemented in detecting child abduction.

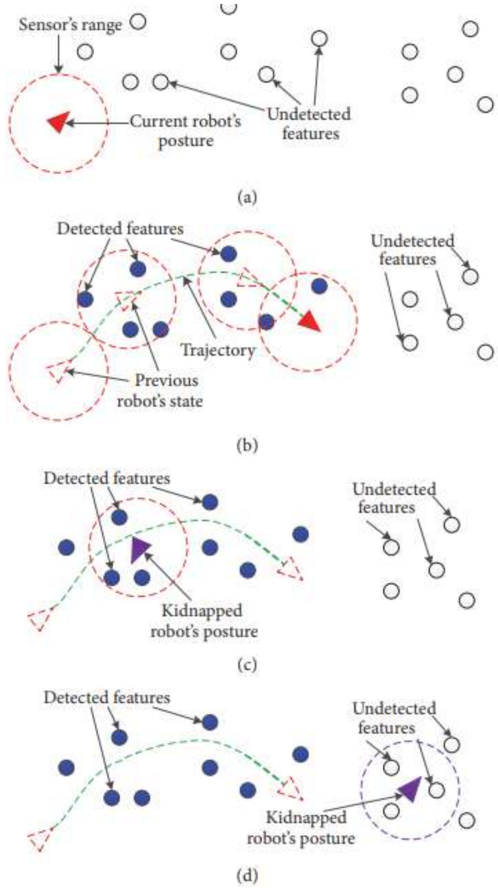


Figure 6: Different situations of kidnapping in simultaneous localization and mapping. (a) Initial situation of simultaneous localization and mapping. (b) Result of simultaneous localization and mapping before kidnapping. (c) The robot is kidnapped to the unexplored area[1]

1.3 Research Gap

Popular researches on kidnap detection has considered the following facts to identify an ongoing kidnap.

- Kidnap by vehicles
- 8Abnormal poses when being kidnapped.
- Frame-based classification for kidnap and normal accompanying.
- Unexpected movements for the target during kidnap

Most of the researches conducted consider abnormal poses as a fact when detecting a kidnap.

All of these available research papers [8] [9] [10] [1] detect kidnaps on activities that happened during the kidnap is taking place and then after proceed to alert or notify the responsible authorities. This does not mitigate the risk of danger of kidnap.[12] states that children face critical mental and physical breakdowns after suffering a kidnap. In such a background allowing a person to be kidnapped for at least a least amount of time cannot be afforded for a child life. Therefore, a common insight these research studies lack is acting upon preventing kidnap.

Child kidnap and overall human or adult kidnap differs in important aspects. Adult kidnapping involves force and might include physical activities like fighting and running. Adults have more knowledge on self-defense than a child. But a child is a mere kid who tends to trust whatever adults say to them [13]. For example, a stranger offering sweets and candies for a child can manipulate the child to follow his instructions and there by will not have to take the kid by force. These serious situations were not considered in the previous researches on kidnap detection. This is noticed as vital gap that should be filled in the future research on child kidnap.

The table 1 shows a tabularized format of the explanation.

Table 1: Comparison of former researches

Research	Detect child kidnaps	Consider common characteristics of a kidnapper for the research	Prevent kidnap before happening	Give alert or notification on detecting a kidnap
[1]	✗	✗	✗	✓
[8]	✗	✓	✗	✓
[9]	✗	✗	✗	✓
[10]	✗	✗	✗	✓
AICare	✓	✓	✓	✓

The AICare concept which is proposed, is designed with many more functionalities than other researches which are currently prevailing. By the proposed solution, the intelligent surveillance system will alert of a possible kidnap if a kidnap suspect is identified within the given criteria and thus prevent the kidnap from happening.

1.4 Research Problem

Nowadays human lives accelerate at an unbelievable speed forcing everyone living in this generation to live with the pressure of engaging in numerous activities within a time frame of 24 hours narrowing down one’s leisure time to a negligible figure. The role of a working mom is no brainer but an obvious victim of the explained situation as her life is sandwiched between the two roles: role of a mom and a wife’s role. Due to the pandemic situation most of the corporates changed their rosters and working plans to work from home. It is given that with the current situation, safety plays a huge role but due to the current working arrangement followed by most offices, mothers are forced to balance office work along with balancing the two roles mentioned earlier which can be hectic.

Given the scenario is such, a babysitter sounds like a good idea but then again leaves us with a question mark on how safe it is. Even a babysitter can make mistakes. Humans are prone to mistakes [14]. Not every person is good at multitasking and it is proved that multitasking is subjected to inefficiency and more mistakes [6]. There arises the problem of safety of children.

According to USA statistics an average of 350 children are abducted every year in USA. Therefore, every day at least one child has been abducted in the world. A child's main responsibility falls on the parents. As per the current evolution of the world not many parents are free enough to pay their hundred percent attention to their child. Children in the pre school age are at the greater risk due to this reason.

In an era where technology develops and families disperse and children solitudes, technology is the only aspect that can bring solutions for these problems.

2. OBJECTIVES

2.1 Main Objective

The main objective is to ensure that children in the early development stages are safe from kidnappings. This safety is achieved by an intelligent surveillance system that is placed in the area where the child is present. The proposed solution will be implemented with the explained motive by indicating the caretakers or parents of such incidents and to prevent or diminish harm to minors as the solution makes sure to notify the relevant parties subsequent to detecting kidnapper suspect.

2.2 Specific Objectives

In order to reach the main objectives, the specific objectives that needs to be attained is as follows,

1. To identify whether a person is suspiciously face covered.

   Determine the most suitable technique that can be used to identify suspicious face coverings accurately and effectively.

2. Authenticate whether a person is authorized to be present in house.

   Use face recognition methodologies to authenticate personnel authorized to be present in the house with the child.

3. Identify suspicious motion speeds of entering persons.

   Determine the most suitable method that can accurately and effectively identify suspicious motion speeds of people entering the house/room accurately and effectively.

4. Identify whether a person is holding sharp objects in hand.

   Discover the most effective object detection methodology that is capable of detecting sharp objects accurately.

3. METHODOLOGY

3.1. System Diagram

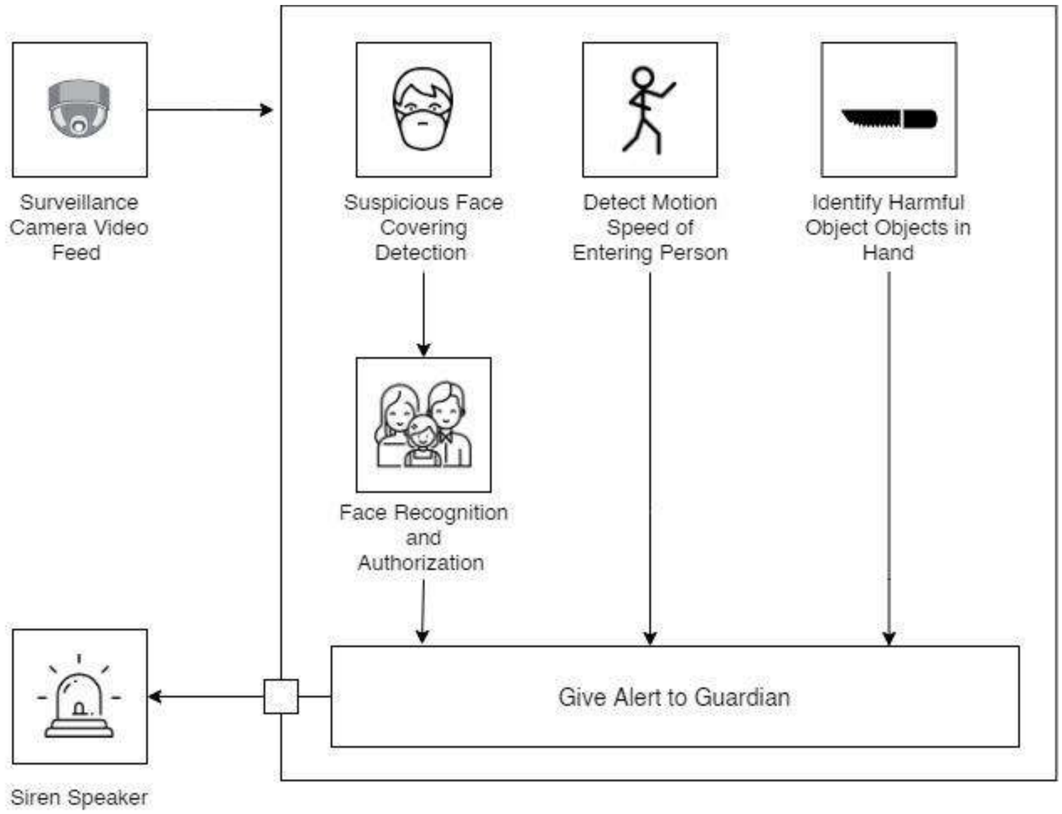


Figure 7: System overview diagram

3.2. Procedure

Identifying a child kidnap is based on detecting kidnapper specific characteristics in a decision tree-based order. The below characteristics of a kidnapper was considered.

1. Face suspiciously covered
2. Unauthorized to enter to house premises
3. Holding shape objects
4. Suspicious motion speed (moving in a hurry)

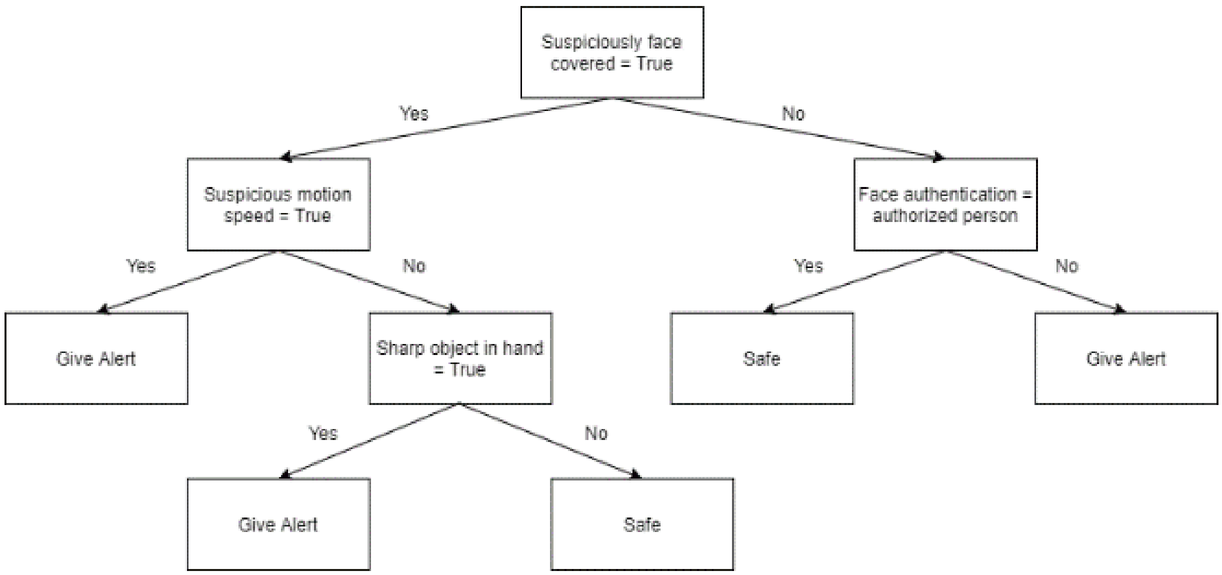


Figure 8: Decision tree to identify a suspect kidnapper

Further explaining the logic of the decision tree, every adult with suspicious face covering, moving a suspicious speed, and holding a sharp object is considered as a kidnapper suspect. Every person who is identified as not authorized to enter is considered a serious case and thus marked as kidnapper suspect. The main purpose of this system is to identify a kidnap before it happens and therefore identifying kidnapper suspects is very essential for the correct operation of the research system.

In the aspect of implementation, the Darknet framework was chosen to develop the application to overcome speed and computation constraints.

Authorized faces for face authentication are stored in a known people data folder. These face images are converted from BGR (OpenCV ordering) to RGB (dlib ordering) to prepare face recognition. Face recognition locates the faces, computes the face embedding, and saves the encodings and the name in a dictionary. The dictionary data is saved to a face encoding file using pickle. In preparation for the encoding file, start person detection in the video stream using YOLOv4.

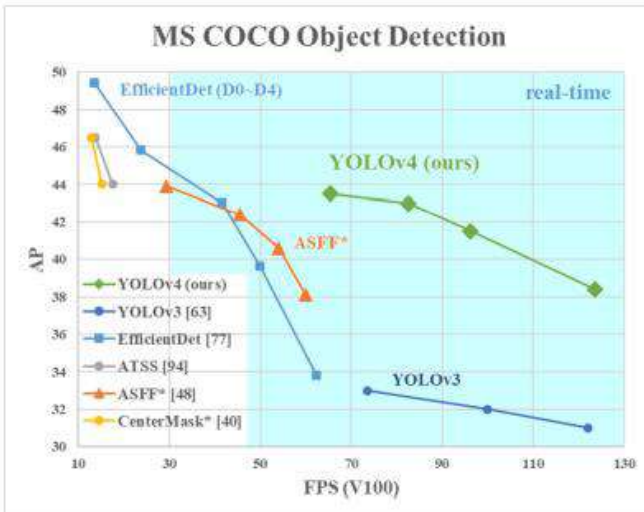


Figure 9: Comparison of YOLOv4 with other state of the art object detectors [15]

When comparing YOLOv4 with other state of the art object detections is observed that YOLOv4 is much better. With comparable performance, YOLOv4 is twice as quick as EfficientDet. YOLOv4 improves the average precision and frame per sedond of YOLOv3 by 10% and 12%, respectively. Due to high performance and accuracy YOLOv4 was thereby chosen to do person detection in this research.

Person detected video frames are then subjected to face detection. The cascades were used to capture the facial features. Out of the two cascades, "haarcacase_frontalface_default" and "haarcascade_frontalface_alt", the former was chosen because identifying as many faces as possible is the main goal at this stage of the research. The cascades pose fewer accuracy issues, but it is eliminated by using face recognition upon the detected faces.

Table 2: Comparing the cascades

	recall	precision	No. of faces detected	No. of true positives
haarcascade_frontalface_default	High	Low	High	Low
haarcascade_frontalface_alt	Low	High	Low	High

Detecting a person and not detecting a face on the frame flags the instance as suspiciously face covered. Person detected, and face detected frames and proceeded for face recognition. Face recognition compares the face embeddings in the face encoding file against the newly computed embeddings of the captured faces in the video frames. Locating matching pairs reject the risk hypothesis and is thus defined as safe, while the opposite prompts an alert indicating the entrance of a possible kidnapper.

Suspiciously face covered instances advance to suspicious motion speed detection. The motion speed is calculated from the distance between the initial and final position of the person bounding box in a defined short period. Speed is calculated horizontally and vertically for a person moving to the side and forward, respectively.

$$elapsed\_time = current\_time - initial\_time$$

$$distance\_horizontal = abs(final\_position\ [0] - initial\_position\ [0])$$

$$distance\_vertical = abs(final\_position\ [1]) - abs(initial\_position[1])$$

$$speed\_sideways = distance\_\ horizontal/elapsed\_time$$

$$speed\_forward = distance\_\ vertical/elapsed\_time$$

The speed is calculated in pixels per second. Once the speed is below the threshold value, it is classified as suspicious motion speed and thereby progress to detect sharp objects in hand. Else prompt an alert indicating the entrance of possible kidnap. The sideways threshold value (4.43) and forward threshold value (1.83) was defined by the average output speed of 6 test videos that logically identified as the average human walking speed of men and women between the ages of 20 - 69 by [16].

Table 3: Calculate average walking speed

Test Case	Result horizontal speed	Result vertical speed	Cumulative avg. horizontal speed	Cumulative avg. vertical speed
Video 1	4.3	0	4.3	0
Video 2	4.0	0	4.15	0
Video 3	4.7	0	**4.43**	0
Video 4	0	1.5	0	1.5
Video 5	0	2.0	0	1.75
Video 6	0	1.9	**0**	**1.83**

Subcomponent of another team member was reused for the sharp object detection portion. Detecting sharp objects in hand was implemented by creating a detector for knives. Transfer learning based methodology was followed considering performance and resource utilization. YOLOv4 and Tiny-YOLOv4 are the two data models trained in the process of developing the detector. To develop the detector, the dataset was obtained by pulling out frames consisting of knives which were collected using multiple sources. Footage of the home surveillance cameras available on the internet, CCTV camera footage captured by ourselves, YouTube videos, and Google images contributed to the image obtaining process, resulting in 3234 images. During preprocessing, the gathered images were converted into a single frame size (1200pixels x 1200pixels) and then were forwarded for annotation. The annotation was done using the LabelImg tool. LabelImg is a graphical image annotation tool written in python. The XML files resulting in the annotation process were converted into the file format compatible with the darknet framework. This conversion was done by using Roboflow Software. Roboflow is a developer tool for building computer vision models.

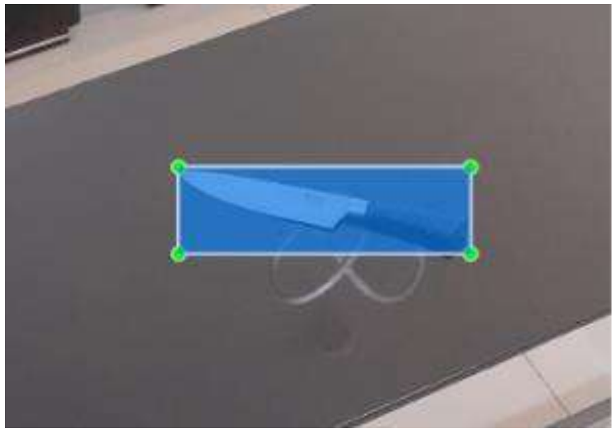


Figure 10: Annotated knife image

Training and testing data were divided as 75% and 25%, respectively, and fed to the YOLOv4 and Tiny-YOLOv4. The purpose of training these two models is to select the most accurate and efficient model for developing the detector.

Table 4: Comparing the models

Model	Avg. Precision	True Positives	False Positives
YOLOv4	97.27%	63	14
Tiny-YOLOv4	99.70%	65	6

With the obtained results, Tiny-YOLOv4 was chosen to be used as the detector for the sharp object.

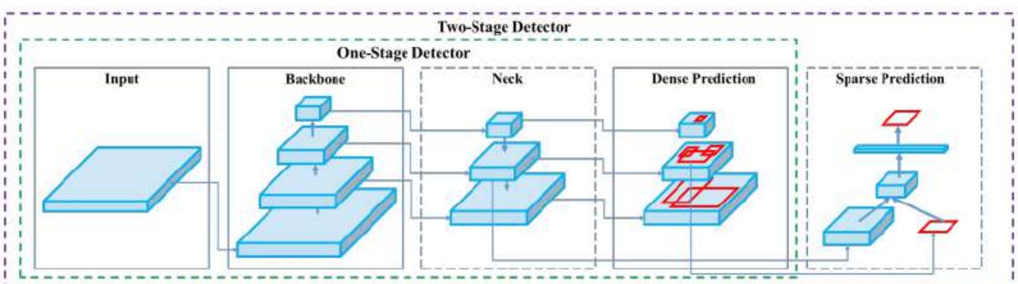
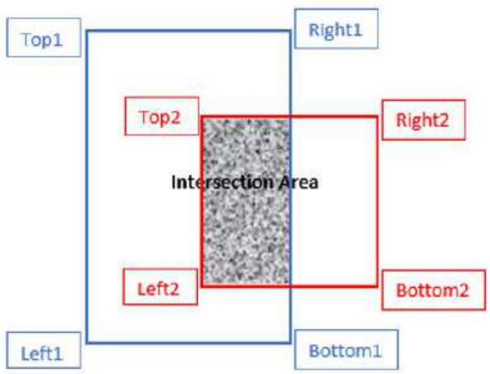


Figure 11: Object Detector of YOLOv4

Elaborating the details of YOLOv4, the backbone of YOLOv4 is CSPDarknet53. The neck of YOLOv4 can be either SPP or PAN and the head which perform dense prediction and sparse prediction is YOLOv3.

Bounding boxes of the person and the object were used to identify whether the person was holding the object. When the bounding boxes overlap, it is considered as the person is holding the knife. This logic is implemented by using a calculation to measure the width and height of the intersection area of the bounding boxes. Resulting in zero width and height means there is no intersection between the bounding boxes of the person and the object. The calculation is demonstrated in the image below.




```
x = max(Left1, Left2)
y = max(Top1, Top2)
width = min(Left1 + Right1, Left2 + Right2) - x
height = min(Top1 + Bottom1, Top2 + Bottom2) - y
```

Figure 12: Calculating the intersection area of the bounding boxes

Explaining the image, bounding boxes are drown on the person and the sharp object detected. An intersection area is created when the person bounding box(blue) and the sharp object bounding box(red) overlap. The area of the intersection is calculated and is the area is above and not equal to zero it is considered that the person is in possess of the relevant sharp object. If sharp objects in hand emanate a positive result, then the hypothesis of occurrence of possible kidnap is amplified and accepted, which leads to prompt alarming. Otherwise, proceed as all safe

3.3 Tools and Technologies

The core of the implementation of this system is based on the concepts of Artificial Intelligence. The domains that were applied in this system are listed below in Table.

Table 5: Applied domains

Domain	Usage
Computer Vision	To perform person and motion speed detection. To detect body positions and movements in order to visually monitor kidnapper characteristics.
Machine Learning	To perform sharp object detection. Train model to detect knife objects of all shapes and sizes. Classification based decision tree to identify kidnapper suspect.
Deep Learning	To perform facial recognition. Extract facial features and compare to known database.

To maximize the usage of the above-mentioned domains, it was essential to use necessary software tools and libraries. The below table displays the list of tools and libraries that was implemented in this system.

Table 6: Tools and libraries

Tool/Library	Usage
Darknet	This is an open-source neural network framework written in C and CUDA that supports CPU and GPU computation.
OpenCV	Library consisting of programming functions mainly aimed at real-time computer vision.
Face Recognition	Library for performing face recognition and manipulation using python which is built using deep learning.
Numpy	To perform mathematical computations in python.
Pickle	To save python object in disk. Used to serializing and deserializing python object structure.
Base64	To encode and decode data.

3.4 Project Requirements

Functional Requirements

- Integration should be allowed between subsystems.
- There should be a way to capture the persons face.
- There should be a way to identify child separately.

Non-Functional Requirements

- Response time and net processing time.
- Efficiency
- Availability

Personal Requirements

- Parent/Guardian should be available.
- Child should listen to the warnings.
- Parent/Guardian should react to the alerts.

Hardware Requirements

- There should be a way to configure the speaker to the system
- There should be a way to configure the camera to the system

3.5 Commercialization

In a society where families with both parents working has become a common norm, child ren has left to grow up by themselves. Children between the age 1 year and 5 year is the most crucial period where a child need a lot of parental attention. Thereby, AICare has the potential to be the latest trend in childcare in the coming decade. Being able to give real time protection assistance to a child when parents are attending to work increase the average working time of an employee. Being able to work from home reduces the number of leaves an employee might take. We anticipate that AICare is going to be a top solution companies will invest on providing for their employees because of the high return of investment AICare provides.

3.5.1 Market potential

The global child safety products market size is anticipated to reach USD 132.2 billion by 2025, according to a new report by Grand View Research, Inc., expanding at a CAGR of 5.0% over the forecast period. Rise in government initiatives regarding infant safety has been driving the global market.

From 2019 to 2025, Asia Pacific is anticipated to witness at the fastest CAGR of 5.8 percent. The market in this region has been primarily driven by an increase in the birth rate and the number of working mothers in countries such as China and India. The female labor participation rate has increased by 4.1 percent over the last three decades, according to EPRA International Journal of Economic and Business Review.

The rise in the number of new births in this region, and also the large population, are the factors driving the growth of the child safety products market in this region. According to the National Health and Family Planning Commission (NHFPC), the average number of new births in China each year from 2016 to 2020 is estimated to be 17 million to 20 million.

3.5.2 Target Market

Working parents with children in the age group of 1 - 5 years old are priority target group of this product. With the current working from home context, parents find it difficult to provide fulltime constant attention to their children even at home. Situations such as when parent is having office meetings while the child is playing in his room are evident proof to the uncomfortable arrangements parents have to face while working from home. The curve of working mothers has rapidly increased during the past century. The high cost of living has made it tough for one parent to support the entire family. With this constraint many mothers has been obligated to engage with an employment in order to co-support the family. At this situation it is foreseeable that AICare will have high demand within working parents.

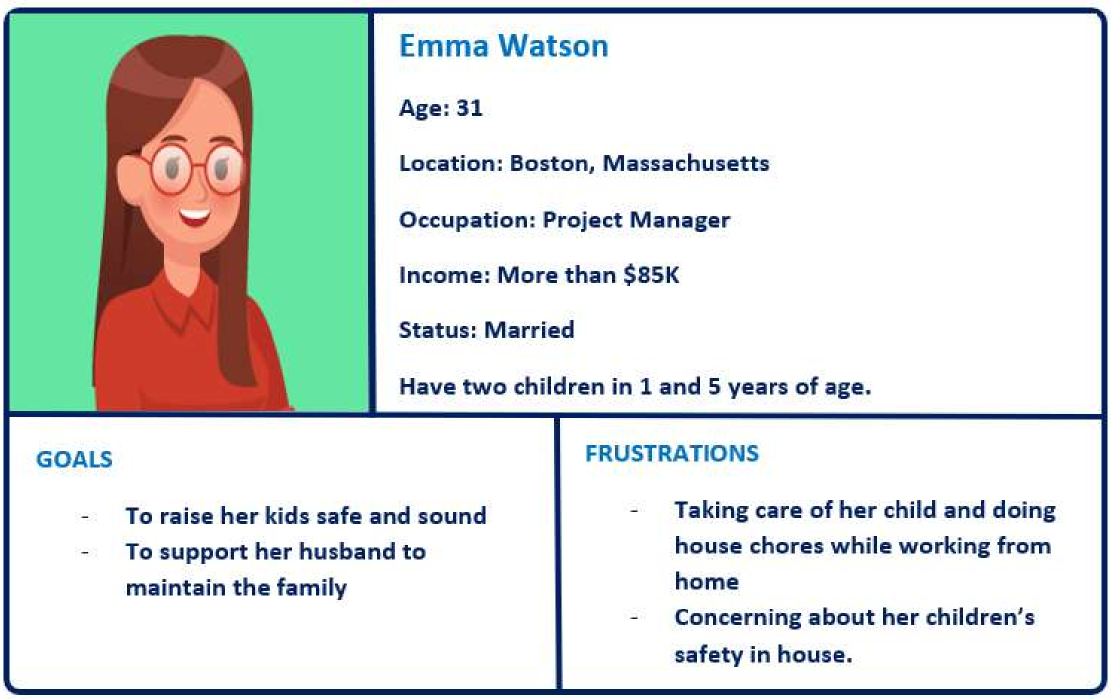


Figure 13: User persona of working from home parent

Companies such as Facebook and Twitter have started to embrace the permanent work from home culture. Thinking of the on-premises cost cutting benefit the companies get, it is expected for more and more companies to embrace forever-work-from-home. Some companies return real-estate savings to staff by offering home office set-up reimbursements or monthly stipends. We anticipate that AICare is going to be one of the top products companies will invest to support work from home parents in the next five years.

JOSH HUTCHINS

Age: 45

Location: London, England

Occupation: Chief Executive Officer

Income: More than $400K

Status: Married

Very dedicated to his company.

GOALS

- **To support work from home employees**
- **Reduce on premise costs of the company**

FRUSTRATIONS

- **Employees less attention to job because of the parental commitments at home when working from home**

Figure 14: User persona of a chief executive officer of a company supporting working from home

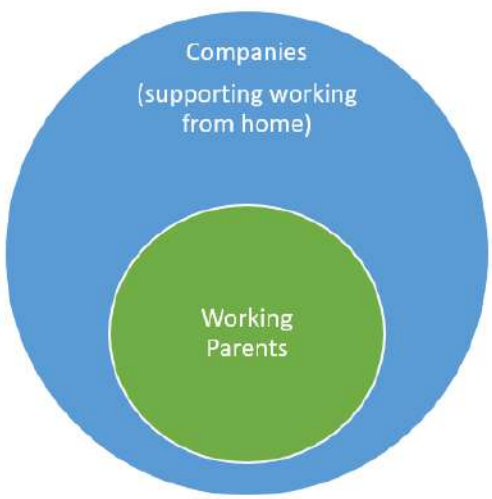


Figure 15: Target market for AICare

3.5.3 Competitive Analysis

In this section, the currently available child safety products in the market will be compared and understood to get a clear picture of the competition and the advantage of demand and supply that exists.

Table 7: Competitor analysis

	Mommy I'm Here, Child Locator	**Amber Alert GPS V3**	**Child Angel**	**Drop Cam HD Video Monitoring System**
Product Specific	Includes a transmitter and a receiver. Transmitter is with mother and receiver can be attached to the baby's shoe or belt. If the baby gets lost, mother can press the button in the transmitter, and it will set an alarm attached to the receiver. Price - $49	A tracker that includes voice recognition and command. Provides the child's location. Allow two-way communication between parent and child with one press of button. Safety zones can be set up and can receive alerts when child leave safety zone. Price - $135	Location monitoring child tracking device. Measure child's daily fitness level. Allows moving geo fence to create safe zone radius around the child. Provide an alarm when the strap is removed. Price - $119	Allow parents to remotely monitor and supervise what is going at home through video monitoring. Price - $177
Target Customer/Message	Parents with toddler kids.	Parents with schoolings children.	Parents with schooling children.	Working parents who leave their child with sitters at home.
Positioning	Only able to locate the child if the parent and child is within 30m. Not a one-time solution, should replace the product time to time. Does not protect the child from injuries or provide preventive measures.	Power lasts only upto 40 hours. Does not protect the child from injuries or provide preventive measures.	Power lasts only upto 48 hours. Does not protect the child from injuries or provide preventive measures.	Does not provide safety measures. Does not protect the child from injuries or provide preventive measures.

3.5.2 Business Model

This product has market potential in the areas of Child Safety and Protection. Adhering to the requirements of the working parents in providing in house protection for their children, we have introduced a unique business model to attract potential clients. Initially, a free plan would be provided to get acquainted with the system. If the customer (Parents/Companies) becomes interested in the product, they can switch for the paid version where the required equipment (Cameras and Speakers) will be installed by charging for the cost of the installation. Interestingly, this is a one-time payment for the user to utilize this product for their needs. A brief summary of the business model is provided in the figure below.

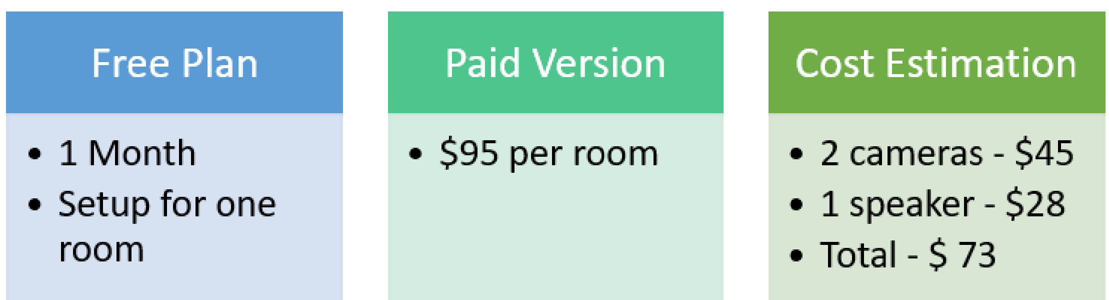


Figure 16: Business model of AICare

4. TESTING AND IMPLEMENTATION

4.1. Implementation

4.1.1. Models

From the collected dataset for sharp objects, some images were manually preprocessed to crop out the irrelevant parts of the image and to focus on the object in the image. When preparing the dataset, the different perspectives of object positions, clearness of the images and lighting conditions were also considered. Once the models were trained, they were saved as .h5 files and used for the sharp object detection in the system.

4.1.2. System

The overall system was implemented by using the Python programming language. The module was implemented by using PyCharm Professional 2021.2.2 provided by Jetbrains Inc. Ltd. The Operating System that supported this implementation was Windows 10 Home. In order to execute the models, a combination of *OpenCV* libraries were required. *Face Recognition* library was used to perform face authentication.

4.2. Testing

4.2.1 Unit Testing

The importance of a well-planned testing procedure is essential to track the existing issues in the components at the very early stages and ultimately reduce the time and cost of the system at the latter stages of development. Hence, unit testing was carried out separately for each main component of the research.

4.2.1.1 Suspicious Face Covering and Face Authentication

In this unit testing, the suspicious face covering and face authentication subcomponent was tested. The main objective here was to identify whether the component is successfully facilitating the initially planned areas. The tests were conducted, prioritizing the following cases.

Case 1: A known person with a suspicious face covering

Case 2: An unknown person with a suspicious face covering

Case 3: A known person without a face covering

Case 4: An unknown person without a face covering

Figure 17: Test results of suspicious face covering and authentication

Evaluating the outcome of the testing, it was observed that the subcomponent is functioning as expected and is capable to capture suspicious face covering occasions of known and unknown peoples and also is able to identify known people faces in proper lighting conditions. However, the component is not capable of capturing suspicious face covering and face authentications in low light conditions.

4.2.1.2 Suspicious Motion Speed

In phase two of unit testing, the suspicious motion speed component integrated with the suspicious face covering and face authentication components was tested. The main objective here was to identify whether the component is successfully facilitating the initially planned areas. A combination of phase one test cases with the below conditions was used.

Case 1: Moving at a suspicious speed
Case 2: Not moving at a suspicious speed

Figure 18: Test results of suspicious motion speed detection

Evaluating the outcome of the testing, it was observed that the subcomponent is functioning as expected and is capable to capture occasions of suspicious motion speed in an inhouse area. However, the component is not capable of capturing suspicious motion speed outside the surveillance camera view sight. To eliminate this limitation to some extent, two surveillance cameras are installed in the room in two locations. Still this will not be sufficient to large rooms with different architectural passage ways.

4.2.1.3 Sharp Objects in Hand

The component of detecting sharp objects in hand was unit tested in phase three of unit testing. The objective of this testing was to check whether the component is successfully facilitating the initially planned areas. The sharp object is expected to be detected as dangerous only if a person holds it. Therefore, the below mentioned vital areas were thoroughly tested.

Case 1: The knife laid bare in the surrounding
Case 2: The knife held by a person

Figure 19: Test results of sharp object detection

Evaluating the outcome of the testing, it was observed that the subcomponent is functioning as expected and is capable to capture occasions of holding sharp objects in hand. Here only knife is considered. However, the component is not capable of capturing sharp objects in low light conditions. It would be beneficial if many objects were considered such as guns that can be possibly held by a kidnapper.

4.2.2 Integration Testing

The consolidation of suspicious face-covering detection, face authentication, suspicious motion speed detection and sharp objects in hand was tested in the integration testing. The priority of this testing was to identify whether the entire system is operating as expected giving the expected results. Tests were performed around identified 16 test cases.

Test Case	Input								Expected Output	Actual Output	Remarks
	Known person	Unknown person	With suspicious face covering	Without face covering	Moving in suspicious speed	Not moving in suspicious speed	Having sharp objects in hand	Not having objects in hand			
1	TRUE		TRUE		TRUE		TRUE		Alert	Alert	Pass
2	TRUE		TRUE		TRUE			TRUE	Alert	Alert	Pass
3	TRUE		TRUE			TRUE	TRUE		Alert	Alert	Pass
4	TRUE		TRUE			TRUE		TRUE	Save image	Save image	Pass
5	TRUE			TRUE	TRUE		TRUE		No output	No output	Pass
6	TRUE			TRUE	TRUE			TRUE	No output	No output	Pass
7	TRUE			TRUE		TRUE	TRUE		No output	No output	Pass
8	TRUE			TRUE		TRUE		TRUE	No output	No output	Pass
9		TRUE	TRUE		TRUE		TRUE		Alert	Alert	Pass
10		TRUE	TRUE		TRUE			TRUE	Alert	Alert	Pass
11		TRUE	TRUE			TRUE	TRUE		Alert	Alert	Pass
12		TRUE	TRUE			TRUE		TRUE	Save image	Save image	Pass
13		TRUE		TRUE	TRUE		TRUE		Alert	Alert	Pass
14		TRUE		TRUE	TRUE			TRUE	Alert	Alert	Pass
15		TRUE		TRUE		TRUE	TRUE		Alert	Alert	Pass
16		TRUE		TRUE		TRUE		TRUE	Alert	Alert	Pass

Figure 20: Test cases covered in the final testing

The above test cases were formed covering all condition combinations that lead to detecting possible child kidnap before it occurs. The yellow cells represent the conditions covered in the specific test cases. Pink and blue columns indicate the expected and actual outputs of the tests, and the green column defines the status of the test. During testing an outlying result occurred that brought into notice about an error in the system which was later corrected. A known person having the face suspiciously covered was identified correctly which is unexpected of the system. This was due to having comparative large number of images in the known faces database from this specific person. Correcting the error, steps were taken to store equal number of images from each known person in the know faces database. After correcting the error, the system functioned as expected.

Figure 21: Test results of overall integration testing

Evaluating the outcome of integration testing, it was observed that the system as a whole is functioning as expected and is capable to capture all possible combinations of suspicious face covering, face authentication, suspicious motion speed and occasions of holding sharp objects in hand for known and unknown persons. However, the system is not capable of capturing possible kidnapper in low light conditions.

5. RESULTS AND DISCUSSIONS

5.1 Suspicious Face Covering and Face Authentication

This section mainly focuses on the discussion and findings of the suspicious face covering and face authentication component. This section provides a summary of findings and figures gathered throughout the testing phase. The explained approach was tested using 8 video clips with the MP4 format, consisting of 4 videos of known person and 4 of unknown person each having two videos separately with face suspicious face covering and without face covering. The results for suspicious face covering are shown in table 8.

- TP = Having suspicious face covering and detecting as having suspicious face covering.
- TN = Not having suspicious face covering and detecting as not having suspicious face covering.
- FP = Not having suspicious face covering but detecting as having suspicious face covering.
- FN = Having suspicious face covering but detecting as not having suspicious face covering.

Table 8: Evaluation of results of suspicious face covering sub component

N = 8	*Actual: suspicious*	*Actual: not suspicious*
detected: suspicious	4 (TP)	2 (FP)
detected: not suspicious	0 (FN)	2 (TN)

The results for face authentication are shown in table 9.

- TP = Known person and detecting as a known person.
- TN = Unknown person and detecting as unknown person.
- FP = Unknown person but detecting as known person.
- FN = Known person but detecting as unknown person.

Table 9: Evaluation of results of face authentication sub component

N = 8	*Actual: known*	*Actual: unknown*
detected: known	2 (TP)	0 (FP)
detected: unknown	2 (FN)	4 (TN)

To evaluate this technique, we utilize a set of criteria that include accuracy, sensitivity, and specificity [17]. figure 22 shows equations of these criteria:

$$\text{accuracy(A)} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{sensitivity} = \frac{TP}{TP + FN}$$

$$\text{specificity} = \frac{TN}{TN + FP}$$

Figure 22: Equations of evaluation criteria[17]

The final suspicious face covering and face authentication component performance is shown in table 10 and table 11 according to the evaluation criteria followed.

Table 10: Performance results of suspicious face covering sub component

criteria	**score**
Accuracy	0.75
sensitivity	1
specificity	0.5

Table 11: Performance results of face authentication sub component

criteria	**score**
Accuracy	0.75
sensitivity	0.5
specificity	1

Here the component was incapable to detect known and unknown persons without face coverings in low light. As the light hinders the features of the person and face the system erroneously interpreted the situation as suspiciously face covered.

5.2 Suspicious Motion Speed

This section focuses on the discussion and findings of the suspicious motion speed component. This section provides a summary of findings and figures gathered throughout the testing phase. The explained approach was tested using 6 video clips (8 sec length) with the MP4 format, which includes videos of suspicious motion speed and non-suspicious motion speed as 3 each. These 3 videos consist of person moving to left side, right side and towards camera. The results for suspicious motion speed are shown in table 12.

- TP = Having suspicious motion speed and detecting as having suspicious motion speed.
- TN = Not having suspicious motion speed and detecting as not having suspicious motion speed.
- FP = Not having suspicious motion speed but detecting as having suspicious motion speed.
- FN = Having suspicious motion speed but detecting as not having suspicious motion speed.

Table 12: Evaluation of results of suspicious motion speed component

N = 6	*Actual: suspicious*	*Actual: not suspicious*
detected: suspicious	2 (TP)	0 (FP)
detected: not suspicious	1 (FN)	3 (TN)

The final suspicious motion speed component performance is shown in table 13 according to the evaluation criteria followed.

Table 13: Performance results of suspicious motion speed component

criteria	**score**
Accuracy	0.83
sensitivity	0.67
specificity	1

Here the component was incapable of detecting person moving towards the camera in a non-suspicious motion speed because in the latter part of the video the person moves too closer to the camera and the system was not able to detect the situation as a person and thus the motion speed calculation was interrupted in the middle.

5.3 Sharp Objects in Hand

This section discusses on the findings of the sharp objects in hand component. This section provides a summary of findings and figures gathered throughout the testing phase. The explained approach was tested using 10 video clips with the MP4 format, which includes videos of knife laid bare and knife held by person as 5 each. The results for sharp objects in hand are shown in table 14.

- TP = Having sharp object in hand and detecting as having sharp object in hand.
- TN = Not having sharp object in hand and detecting as not having sharp object in hand.
- FP = Not having sharp object in hand but detecting as having sharp object in hand.
- FN = Having sharp object in hand but detecting as not having sharp object in hand.

Table 14: Evaluation of results of sharp object in hand component

N = 10	*Actual: knife in hand*	*Actual: knife not in hand..*
detected: knife in hand	4 (TP)	0 (FP)
detected: knife not in hand	1 (FN)	5 (TN)

The final sharp objects in hand component performance is shown in table 15 according to the evaluation criteria followed.

Table 15: Performance results of suspicious motion speed component

criteria	**score**
Accuracy	0.9
sensitivity	0.8
specificity	1

Here the component was incapable of detecting the knife in low light condition. Even though the person was detected and the person’s bounding box overcame the knife in the video, the system was able to capture the knife object.

5.4 Kidnap Detection and Prevention Assistance

Here it is focused on the discussion and findings of the final system after successfully integrating all the components together. This segment provides the summary of findings and figures gathered throughout the integrated testing phase.
The integrated system was tested using 16 video clips with the MP4 format, which represents all possible combination of test cases of the components suspicious face covering, face authentication, suspicious motion speed and sharp objects in hand as explained in table 4. The results for system integration testing are shown in table 16.

- TP = Having a possibility for kidnap to occur and detecting as having a possibility for kidnap to occur.
- TN = Not having a possibility for kidnap to occur and detecting as not having a possibility for kidnap to occur.
- FP = Not having a possibility for kidnap to occur but detecting as having a possibility for kidnap to occur.
- FN = Having a possibility for kidnap to occur but detecting as not having a possibility for kidnap to occur.

Table 16: Evaluation of results of kidnap detection and prevention assistance integrated system

N = 6	*Actual: possible kidnap*	*Actual: no possible kidnap*
detected: possible kidnap	10 (TP)	0 (FP)
detected: no possible kidnap	0 (FN)	6 (TN)

The final kidnap detection and prevention assistance integrated system performance is shown in table 17 according to the evaluation criteria followed.

Table 17: Performance results of kidnap detection and prevention assistance integrated system

criteria	**score**
Accuracy	1
sensitivity	1
specificity	1

By looking at the results, it can be concluded that the solution has a 10% error rate at its maximum, which results in a 90% accuracy. Even though the test results give an accuracy of around 100%, based on the angle the person is visible to the camera, the limited visible area captured to the camera and external conditions like lighting can result in a lifted error rate, which accounts for a 10% error. When comparing the time spent and the accuracy level achieved, we can consider the solution as successful.

6. CONCLUSION

This paper proposes and develops a solution of kidnapper characteristics-based kidnap detection to distinguish the problem definition of real-time child-specific safety assurance in house borders. The main drawback with the current solutions being unable to enter the market is the futile attempt to implement adult safety assurance methodologies in building child safety solutions that do not possess the required capabilities to diagnose child-specific dispositions of harmful activities.

This solution will assist the working from home parents in ensuring the child's safety while attending office meetings and other related work. A real-time kidnap detection and prevention assistant is an application that recognizes a possible kidnap before occurrence and provides immediate preventive measures to avert harm.

The solution leads to a 90% accuracy with an ability to identify possible danger specific to the child differentiating from adults in a defined indoor area real-time and provide prompt alert/warning to prevent probable harm. Therefore, the proposed system has greater capability in assuring child safety and preventing danger addressing the market gap of similar products.

For future works, this research suggests exploring scenarios where a kidnapper might possess a gun. Gun detection technologies[18] can be examined and implemented. This research does not address child kidnap detection in different light conditions. Capturing child kidnap through surveillance cameras has multiple limitations. Can use the currently pertaining human activity recognition technologies[19] to research further about this in the future.

7. REFERENCE

[1] Y. Tian and S. Ma, “Kidnapping Detection and Recognition in Previous Unknown Environment,” *J. Sensors*, vol. 2017, 2017, doi: 10.1155/2017/6468427.

[2] 2 Norsuzila Ya’acob1, Mardina Abdullah1, 2 and Mahamod Ismail1 *et al.*, “We are IntechOpen , the world ’ s leading publisher of Open Access books Built by scientists , for scientists TOP 1 %,” *Intech*, vol. 32, pp. 137–144, 1989, [Online]. Available: http://www.intechopen.com/books/trends-in-telecommunications-technologies/gps-total-electron-content-tec- prediction-at-ionosphere-layer-over-the-equatorial-region%0AInTec.

[3] A. . News, “Abduction of 2 kids from NY foster home,” 2021.

[4] J. Saranya and J. Selvakumar, “Implementation of children tracking system on android mobile terminals,” *Int. Conf. Commun. Signal Process. ICCSP 2013 - Proc.*, pp. 961–965, 2013, doi: 10.1109/iccsp.2013.6577199.

[5] Y. Mori *et al.*, “A self-configurable new generation children tracking system based on mobile ad hoc networks consisting of android mobile terminals,” *Proc. - 2011 10th Int. Symp. Auton. Decentralized Syst. ISADS 2011*, pp. 339–342, 2011, doi: 10.1109/ISADS.2011.51.

[6] S. C, “Guardian Uses Bluetooth Low Energy Tech to Keep Child Safe.”

[7] G. Schumacher, *Fingerprint Recognition for Children*. 2013.

[8] A. Miyahara and I. Nagayama, “An intelligent security camera system for kidnapping detection,” *J. Adv. Comput. Intell. Intell. Informatics*, vol. 17, no. 5, pp. 746–752, 2013, doi: 10.20965/jaciii.2013.p0746.

[9] J. H. Park, K. Song, and Y. Kim, “A Kidnapping Detection Using Human Pose Estimation in Intelligent Video Surveillance Systems,” *J. Korea Soc. Comput. Inf.*, vol. 23, no. 8, pp. 9–16, 2018, doi: 10.9708/jksci.2018.23.08.009.

[10] R. H. Gwon, K. Y. Kim, J. T. Park, H. Kim, and Y. S. Kim, “A kidnapping detection scheme using frame-based classification for intelligent video surveillance,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8170 LNAI, pp. 345–354, 2013, doi: 10.1007/978-3-642-41218-9_37.

[11] A. Sofranova, “12 Signs that can help you Recognize a child kidnapper,” *Bright Side*.

[12] M. Freeman, “International Child Abduction : Research on the Effects of Abduction and Reunification THE EFFECTS OF CHILD ABDUCTION AND REUNIFICATION - TODAY ’ S PRESENTATION,” no. June, 2019.

[13] N. Branan, “Reviews and Recommendations,” *SA Mind*, 2009.

[14] CogniFit, “Divided Attention,” 2021.

[15] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” no. May, 2020, [Online]. Available: http://arxiv.org/abs/2004.10934.

[16] R. W. Bohannon and A. Williams Andrews, “Normal walking speed: A descriptive meta-analysis,” *Physiotherapy*, vol. 97, no. 3, pp. 182–189, 2011, doi: 10.1016/j.physio.2010.12.004.

[17] S. G. Miaou, P. H. Sung, and C. Y. Huang, “A customized human fall detection system using omni-camera images and personal information,” *Conf. Proc. - 1st Transdiscipl. Conf. Distrib. Diagnosis Home Heal. D2H2 2006*, vol. 2006, pp. 39–42, 2006, doi: 10.1109/DDHH.2006.1624792.

[18] S. Loganathan, G. Kariyawasam, and P. Sumathipala, “Suspicious Activity Detection in Surveillance Footage,” *2019 Int. Conf. Electr. Comput. Technol. Appl. ICECTA 2019*, pp.

1–4, 2019, doi: 10.1109/ICECTA48151.2019.8959600.

[19] K. G. Manosha Chathuramali and R. Rodrigo, “Faster human activity recognition with SVM,” *Int. Conf. Adv. ICT Emerg. Reg. ICTer 2012 - Conf. Proc.*, pp. 197–203, 2012, doi: 10.1109/ICTer.2012.6421415.

8. APPENDICES

Appendix A: Turnitin Similarity Report

ORIGINALITY REPORT

16%	13%	8%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Sri Lanka Institute of Information Technology Student Paper	4%
2	ijarcce.com Internet Source	4%
3	www.grandviewresearch.com Internet Source	2%
4	www.hindawi.com Internet Source	1%
5	www.springerprofessional.de Internet Source	1%
6	"An Intelligent Security Camera System for Kidnapping Detection", Journal of Advanced Computational Intelligence and Intelligent Informatics, 2013 Publication	1%
7	www.koreascience.or.kr Internet Source	1%

8	S. N. David Chua, S. F. Lim, S. N. Lai, T. K. Chang. "Development of a Child Detection System with Artificial Intelligence Using Object Detection Method", Journal of Electrical Engineering & Technology, 2019 Publication	1 %
9	www.gensler.com Internet Source	<1 %
10	hackernoon.com Internet Source	<1 %
11	www.irjet.net Internet Source	<1 %
12	ieeexplore.ieee.org Internet Source	<1 %
13	journals.theired.org Internet Source	<1 %
14	www.smgov.net Internet Source	<1 %
15	Chen Chen, Nver Ren, Maoying Li, Xu Cheng. "A Vehicle Trajectory Extraction Method for Traffic Simulating Modeling", E3S Web of Conferences, 2021 Publication	<1 %
16	rdrr.io Internet Source	<1 %

17	www.stemina.com Internet Source	<1%
18	Qingbang Shi, Jun Li. "Objects Detection of UAV for Anti-UAV Based on YOLOv4", 2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT, 2020 Publication	<1%
19	Zhihao Cao, Mingfeng Shao, Li Xu, Shaomin Mu, Hongchun Qu. "MaskHunter: real - time object detection of face masks during the COVID - 19 pandemic", IET Image Processing, 2021 Publication	<1%
20	Al-Lawati, Anwaar, Shaikha Al-Jahdhami, Asma Al-Belushi, Dalal Al-Adawi, Medhat Awadalla, and Dawood Al-Abri. "RFID-based system for school children transportation safety enhancement", 2015 IEEE 8th GCC Conference & Exhibition, 2015. Publication	<1%
21	Yang Tian, Shugen Ma. "Kidnapping Detection and Recognition in Previous Unknown Environment", Journal of Sensors, 2017 Publication	<1%