HARMFUL CHILD ACTIVITY DETECTION AND PREVENTION ASSISTANCE SYSTEM

2021-115

Final Report

Pramodya Hettiarachchi

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

HARMFUL CHILD ACTIVITY DETECTION AND PREVENTION ASSISTANCE SYSTEM

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

I do hereby declare that this is my work, and this final report does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any other university or institute of higher learning, and to the best of my knowledge and belief, it does not contain any material previously published or written by another person except where the acknowledgment is made in the text.

Name	Student ID	Signature
Pramodya Hettiarachchi	IT18006308	

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

The above candidates are researching the undergraduate Dissertation under my supervision.

……………………………   …………………..

Signature of the supervisor: Date

ABSTRACT

This paper is based on a household hazardous objects detection and prevention detector to ensure child safety at home. The increasing number of accidents in the home causes more than 37% of death in children between the age of one and four due to harmful child activities from hazardous objects like fork, blades, knives, and hot liquid containers. In most of other fields, some approaches that have been proposed over the last few years, that are based on deep learning, have been proved to perform better than the machine learning methods, with the use of computer vision being the most eminent case. I propose a computer vision framework using a deep learning algorithm to identify hazardous objects and prevent dangers in real-time using a surveillance camera, to perform this application. The advancement of artificial intelligence (AI) technology, especially in image processing and object detection, could be applied to overcome the bottlenecks of the existing devices of harmful activity detection, that often fail to serve as an alerting system for guardians to protect their children from household accidents. The purpose behind developing a proper mechanism for hazardous object detection and prevention, is to enhance and strengthen the existing child security system. The key is to prevent the child accident in real-time, which follows detection of hazardous object and child, identification of proximity range, and proper triggering of alerts. Object detection is used to identify possible hazardous objects in the kid's environment, ie knife, scissor, gas kettle, teacup. Identifying the proximity range is helpful in checking whether the child is in proximity to the hazardous object. Finally, if a child is in proximity to the hazardous object, the alerting system will be triggered accordingly.

The system results in an overall accuracy that is above 90%. As this solution is child-sensitive and responsive, it provides a long-term platform that can monitor potential domestic accidents in real-time based on object detection and proximity theories to support working from home parents to take care of their children during crucial times.

Keywords: Computer vision, Deep Learning algorithm, Hazardous object detection

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) and especially the head of the Research Project module, Dr. Janaka Wijekoon for giving an opportunity to carry out this research project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVATIONS

Abbreviation	Description
CNN	Convolutional neural network
YOLO	You Only Look Once

1. INTRODUCTION

All around the world, it's the normal practice to keep children often at home, for their own safety. Even so, there is always a risk of them getting unintentionally injured, as they cannot identify dangers around them. The organization 'Safe Kids Worldwide' indicates that more than 2,200 children die from domestic injuries yearly. WHO World health statistics show that 21% of child deaths are caused by unintentional injuries in the European region alone. For a child below age 4, almost everything around them is dangerous. For a child between ages 5 - 10, even if they can prevent falls and communicate dangers, some domestic objects can be a cause of injury if not handled with caution.

In this study, I would like to direct your attention towards these domestic hazardous objects and my proposed solution to establish a safe environment for children, preventing injuries from hazardous objects.

The definition of a 'hazardous object' depends on the child's age and development and the environment they live in. In general, it can be any object that would potentially threaten the person handling it, if not handled properly. It is tough to keep a toddler away from such objects, and it is equally hard to create an ultimately 'child-proof' home by removing all the hazardous objects from the reach of the child. It requires dedicated attention to keep a child always safe. With the current pandemic situation, all the working criteria have been limited to the households making the parents are forced to balance the office work along with their kids.

There are many 'child-proofing' methods and devices used in homes in the present. Some can be elaborated as follows.

a) Using corner and edge bumpers

Sharp edges and corners of household items can be covered with bumpers to prevent children from getting harmed by falling over them. However, these may require to be checked frequently as they may become loosened over time. Even if this may be effective with small children, it may not have the same effect on toddlers who can tear them apart.

b) Safety gates

The kitchen or other workspaces can be installed with a child gate at the entrance to prevent babies from going in and handling dangerous objects. However, the gate itself might become dangerous to a toddler if they tried to climb it.

c) Baby monitor

A more advanced method used in the present is installing surveillance cameras in baby rooms for monitoring purposes. This enables parents or caregivers to keep an eye on the child even from a distance. However, the effectiveness of this method is questionable as it does not really differ from parents requiring to always keep an eye on the baby, through a camera or not. The child is still in possible danger when the parent cannot monitor.

The WHO statistics from the European region shows, the death and injury rate among children has not decreased in the past 20 years, which confirms the lack of an effective mechanism to prevent such injuries. Safety gates, bumpers or monitors do not provide parents or authorities with real-time danger alerts if the child somehow encounters one. These methods, as it seems, provide only danger avoidance, not effective prevention.

As a solution to the previously identified problem, I propose incorporating the assistance of an Artificial Intelligence-based, hazardous object identifying and alerting system, which would use real-time camera footage to detect dangerous objects in an area.

The proposed solution is capable of

- Detecting hazardous objects using a live video feed.
- Identifying if the child is in proximity to a hazardous object.
- Identifying the danger level of the hazardous object.
- Trigger the appropriate warnings

Parents can determine the safest distance at which they would receive a danger alert and which objects to do so. When they get an alert, they can immediately attend to the problem and prevent any unfortunate event from happening

As stated previously, it is common knowledge that preventing child injuries can be problematic, specially when children begin to walk and explore the environment. Activities like walking, feeling and handling things are needed for a child to develop physically as well as mentally. We can not contain children just for the sake of keeping them safe.

As such, it is very important to adapt a proper procedure in which both the children can be safe and parents can attend to other matters freely

1.1. Background and Literature Survey

When the toddlers start off to learn how to walk, generally between 9 and 16 months, the danger around them increases resulting uncountable tragedies. The number of deaths of kids up to 14 years old is caused by accidental injuries that happened at home even though they are kept at home to protect them from dangers. The main reason for injuries and deaths which are caused by hazardous objects is that the kids do not have the sense of identifying the objects which will be hazardous to them. It is an essential learning process of the development stage of the kids to reach out the things that gain them a curiosity. The more curiosity they gain, the more danger will affect them. The majority of child and youth injuries are preventable [1]. It is a leading health problem and a significant public health issue when considering the frequency, severity and the life-long disability and deaths make injuries. [2]

Parenting takes dedication and time, especially when taking care of the toddlers who are not familiar with the environment and very close to the accidents. However, this may not be practical for most of the parents as they must take care of an actual job to make a living. Thus, a home security camera with baby monitors and IoT based smart baby monitoring system [3] [4] can help parents have the reassurance with a 24/7 view of the child.

However, in recent years, childcare has become more challenging for working parents. Even at home, mothers will not have enough time to monitor their babies continuously, as they have to attend to other essential housework too. Furthermore, in the current context, the evaluation of human civilization and the changing patterns of life along with the current economy, both parents in most of the families have to do a job to ensure that their children are set up for a quality life in the future. Therefore, childcare is more challenging than ever. With these considerations, the previously stated baby monitors and IoT based smart monitoring systems might not be the best answer.

Intelligent Video Surveillance System [5], [6] used to monitor human behaviors, movements, and information which for controlling, managing, and protecting people from danger levels. With the immense support of the technologies and concepts, the usage of automated surveillance has been taking into another level of growth during the last few decades. Following are some of the many areas of surveillance which has been introduced with deep learning and intelligent surveillance concepts.

a. Object Discrimination
b. Action recognition
c. Object and group tracking
d. Object detection
e. Frame-based detection
f. Feature detection and temporal filtering

In the sectors such as crime detection [7], [8]; elder care [9] [10]; accident detection [11], traffic oversee and control [12]; counting moving object [13]–[14]; human activity understanding [15]; motion detection [6], and identifying, tracking, and classifying vehicles, human, and any object of interest [16] can be clearly identify the implementational advancements in video surveillance systems. The utilization of these systems has improved by the increased availability, development and moderate prices of processors and sensors in both indoor and outdoor environments such as shopping malls, airports, train stations, and parking lots [16], [17].

The above-mentioned methods cannot perform using only the video surveillance system. With the time, the complexity nature of the process detection increased making the computational power insufficient to perform the detections. The methodologies such as application-oriented research that integrates computer vision [18], machine learning and image processing arose with the advancements of technology. The primary purpose of this approach is to take a scene in a video as input, automatically interpret it and observe, and predict the interaction of an object in the scene based on information gathered from camera sensors.

At present, Real-time object detection [19] is a viral topic when considering challenges of the applications of the concepts of computer vision. Furthermore, most solutions aim to enhance real-time object detection with the input of video surveillance camera footages using OpenCV and deep learning algorithms. Open CV libraries provide the facility to program for different results under various conditions and thereby assist in finding the optimal threshold and is providential to the edge of the image.

Most researchers implemented computer vision mechanisms for object detection using raspberry pi single-board computers. In most scenarios, OpenCV libraries and Canny edge detection, dilation, and erosion algorithms are used to detect objects. In my research, hazardous object detection and triggering proper alerts to prevent danger in real time are essential when considering child accidents. Nevertheless, the low-performance computing system can only classify objects with acceptable frames rate and accuracy. Therefore, object detection [20] using raspberry pi may not be convenient for children's accidents.

The computer learns how to solve problems without using prewritten algorithms is the idea that based on the Machine Learning which is a branch of artificial intelligence. Predict the result based on the input data is the objective of incorporating machine learning. The flow of traditional machine learning is that a machine is provided with training data and let it to make correct decisions when faced with actual data. In other words, in traditional machine learning, a machine does not have the capability of performing and solving a significant number of tasks unless with the human control.

Deep Learning [21] [22] is a class of machine learning algorithms, resorts to a multi-layered filter system where each successive layer of the input receives the output data of the previous layer, in simple terms the features of a higher level are derived from the features of a lower level. The deep learning models have the capability of making the new features by themselves while in the traditional machine learning approach's features need to be identified accurately by user. As for the key advantages of using deep learning, high accuracy, advance analysis capabilities, adaptability and scalability can be mentioned.

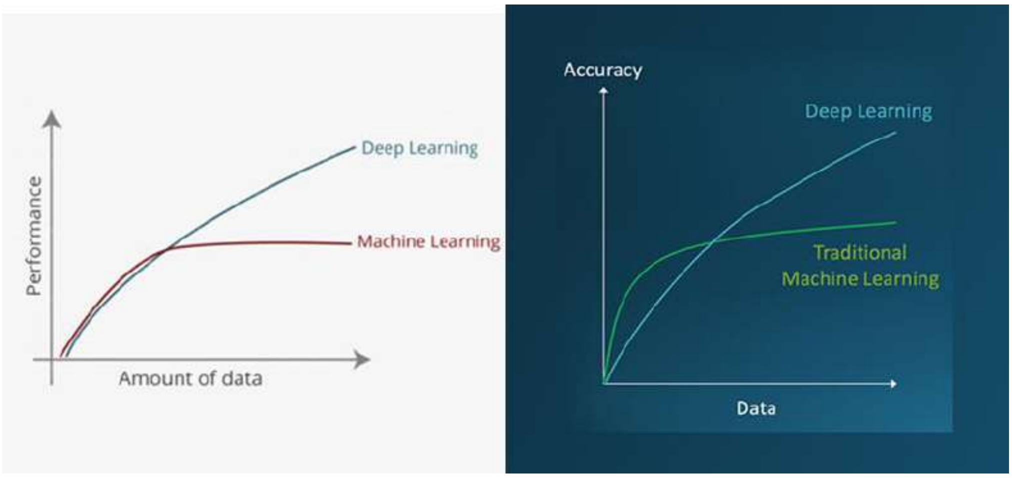

fig 1. Performance and Accuracy Diagram

The utmost challenge in object detection is obtaining specific geometric features from objects. Canny, Hough transform [23], Sobel, Prewitt, and chain codes are some of the existing pattern description methods which do not offer specific features for object recognition along with that. However, these detection methodologies have some drawbacks. When we consider time, the canny detection algorithm is time consuming because of the complex computation, all the above-mentioned algorithms do not provide a feature for object detection with real-time responsiveness.

When considering the process of image classification and object detection, it must be made explicit that these are two different tasks. In classification, as the car in the given example, the classifier outputs the class of a higher-level that the image belongs with a high probability. Whereas in object detection, as the four boxes in the example the detector outputs the bounding box coordinates that localize the detected objects and their predicted classes (cars, trucks, traffic lights, bicycle, person).

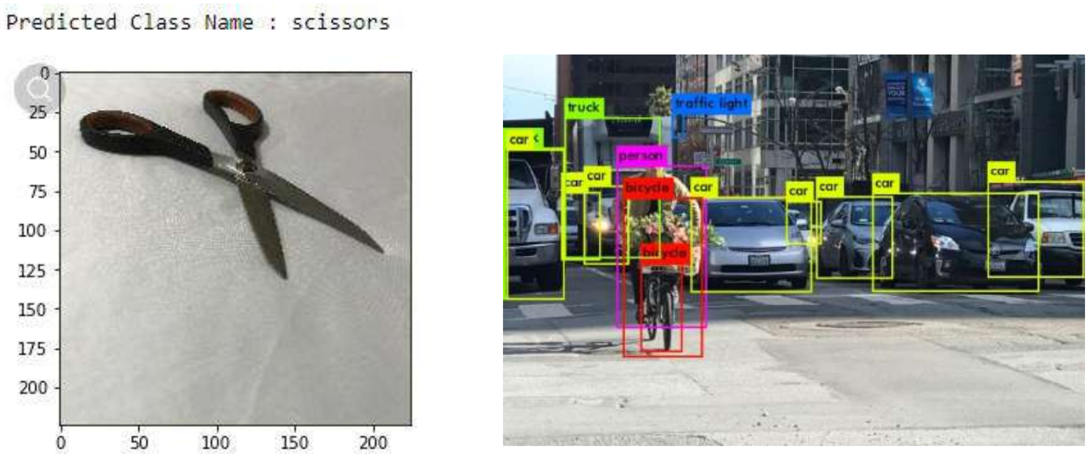

fig 2. Object Classification vs Object Detection

In this research, I have purposed to use object detection task to identify the hazardous objects.

Since the machines can not distinguish the hazardous objects instantly in a video stream, and in real time it is required for the deep learning algorithms to be accurate and fast to detect an object, choosing the right algorithm [24] for detecting an object is very important. As the most efficient algorithms for detecting objects in real-time Faster R-CNN , YOLO [25] and SSD have been used for many approaches.

a. Faster R-CNN

Ross Girshick has introduced the Fast R-CNN model as a method of object detection [26]. As shown in the figure, Faster R-CNN has integrated feature extraction, proposal extraction, rectification in a network. The overall performance and the quality of the proposed box are improved. In terms of detection speed, the number of proposed frames has reduced from about 2000 to about 300.

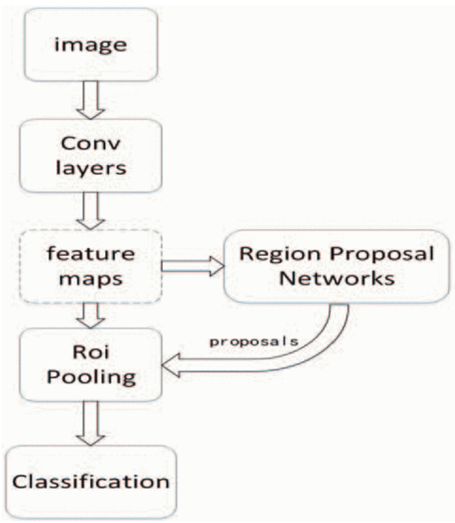

fig 3. Simple Architecture of Faster R-CNN

The four main contents of the structure of Faster R-CNN are as below [27]:

- Convolution Layer
- Region Proposal network (RPN)
- RoI pooling
- Classification

Since Faster R-CNN is a region-based model, the detection happens in two stages, thus called two-stage detectors. The following problems with the above networks were identified since these architectures were not able to manage to create a real-time object detector:

a. Training the data is unwieldy and too long
b. Training happens in multiple phases (e.g., training region proposal vs classifier)
c. Network is too slow at inference time

One-stage detection [28] can be the best solution to overcome the bottleneck of this model. The detectors skip the region proposal stage and run detection directly over a dense sampling of possible locations. This is faster and simpler but might potentially drag down the performance a bit. SSD and Yolo will explain about one-stage object detector.

b. SSD

The Single Shot MultiBox Detector SSD [29] [30], uses a single deep neural network to detect objects in images. SSD is a simple method and requires an object proposal as it eliminates the process of generating a proposal. Along with that , combining everything into a single step, it eliminates the subsequent pixel and resampling stages. In addition, SSD is comparatively easy to train and straightforward quality when integrating it into the system makes object detection easier. The primary feature of SSD is using multiscale convolutional bounding box outputs that are attached to several feature maps.

The architecture of the SSD model is composed of three main parts:

1. Base network to extract feature maps
2. Multi-scale feature layers
3. Non-maximum suppression

When considering about the hazardous object detection, , SSD does worse for smaller objects than bigger objects since Shallow layers may not generate enough high-level features to do prediction for small objects [31]. SSD requires a large amount of data to train. The model described in this paper is pretrained on big datasets such as Pascal VOC, COCO and Open Images before training it on our own data.

c. YOLO

YOLO [25] involves one trained neuron. It takes an image as input and gives a prediction of a bounding box and the class labels as output.YOLO involves one trained neuron. As an input it takes an image and as output it gives a prediction of a bounding box and the class labels.

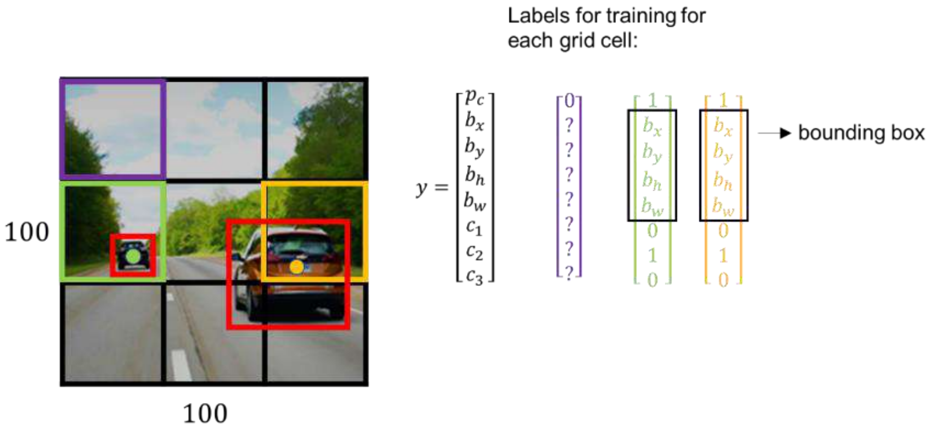

fig 4. YOLO - Imput Image is Spit into grids

YOLO uses an entirely different approach. It applies a single neural network to the full image. This network divides the image into regions and predicts bounding boxes and probabilities for each region. These bounding boxes are weighted by the predicted probabilities. The whole detection pipeline is a single network

In YOLO the detection pipeline is a single network making its approach entirely different where it applies a single neural network to the full image and divides the image into regions and predicts the bounding boxes and probabilities for each divided region.

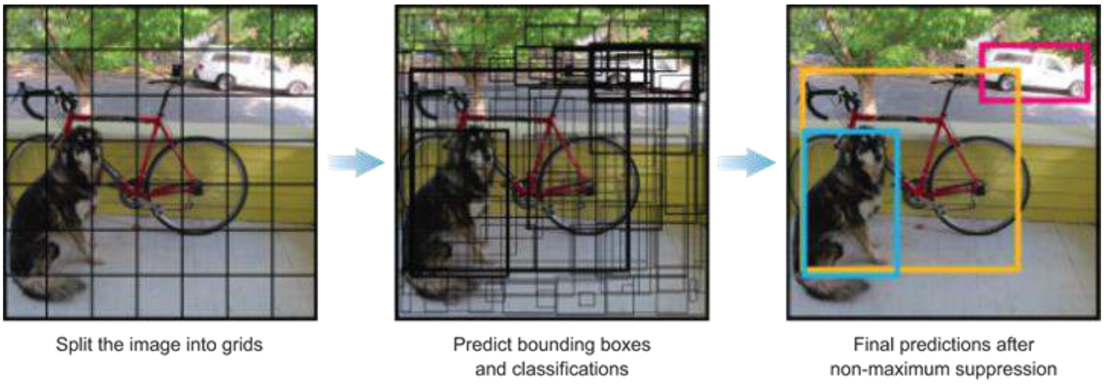

fig 5. YOLO splits the image into grids, predicts objects for each grid, then use non-maximum suppression to finalize predictions

Image Source:

https://colab.research.google.com/github/sony/nnablaexamples/blob/master/interactive-demos/yolov2.ipynb

There are many variants of the YOLO, which have been developed by researchers. But in this paper is discussed only the latest versions of YOLO architectures, Yolov3 and the Yolov4 object detection architectures. Choosing the correct algorithm for the research is very important since this research is based on accuracy and detection speed. The Yolov3 and Yolov4 algorithms are both excellent at object detection. Below are the results obtained using yolov3 and yolov4 on the coco dataset for object detection.

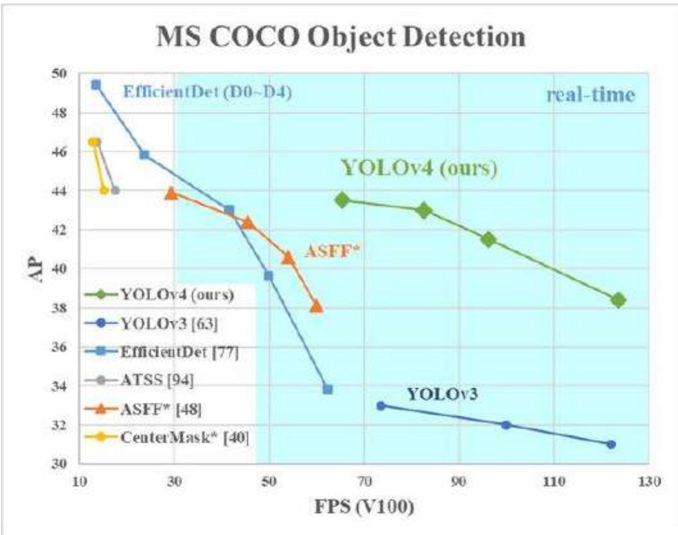

fig 6. .Comparison of the proposed YOLOv4 and other state-of-the-art object detectors

Image source: Source: https://manivannan-ai.medium.com/yolo-v4-750cd627064f

In experiments, YOLOv4 obtained an AP value of 43.5 per cent on the MS COCO dataset and achieved a real-time speed of ∼65 FPS on the Tesla V100, beating the fastest and most accurate detectors in terms of both speed and accuracy.

1.2. Research Gap

Accidents are one of the biggest causes of death in the world. When thinking of accidents, traffic accidents, and outdoor activities-related accidents are the most significant. But accidents happening at home which is regarded as the safest place is less prominent. Children in the early stages of development are moving towards deaths and preventable disabilities as they tend to be exposed to the risk of accidents in the home. When considering hazardous objects accidents, Kids do not have the sense to distinguish hot from cold or sharp from dull. There is a high risk of having danger through hazardous objects to kids life.

To address those child accidents, different approaches have been proposed recently. One of the popular options is a home security camera with a baby monitors system. The drawback of this system is that if parents involve another work, the only way to identify the danger is the baby crying sound. If she could not hear the baby sound, she could not identify the dangerous incidents.

Apart from the baby monitor systems, another solution is a conventional CCTV surveillance system installed in the home primarily used for monitoring the children's activity, tackling emergencies, and ensuring the safety and well-being of the children within the setting. However, only the CCTV surveillance system is not sufficient for child safety. Because, if some accident happens to a child from using the hazardous object, the damage cannot be recovered. As research empowered more influence in computer vision to prevent that accident and damage can be minimized. Detecting objects and identify the danger level has been a challenging task as the objects were diverse. This paper is proposed identifying objects based on sharp objects and hot liquid container objects. Following this way, This system would be able to classify what the object is and alert the parent about the danger of this object. As we know, hot liquid container accident damages are high rather than sharp object accident damages. Only identifying hazardous objects is not enough. This solution also proposes classifying whether a Child is in proximity to a hazardous object or not over the live surveillance footage.

Based on the literature review explanation, very few object detection algorithms work perfectly in real-time. The proposed solution, hazardous object detection speed, is critical compared to the other real-time detection system because the main objective of this system is to prevent the child from danger or minimized accident damage. Furthermore, none of the previous solutions is not explicitly implemented targeting a child.

The research gap details are mentioned given in the below figure.

Product	Research A [2]	Research B [8]	Research C [5]	Research D [3] [4] [5] [6]	**Proposed solution (AI Care)**
Child Surveillance System	✓	✗	✗	✗	✓
Real Time Object Detection	✗	✗	✓	✓	✓
Real Time Sharp Object Detection	✗	✗	✓	✗	✓
Classify the danger level	✗	✗	✗	✗	✓
Child detection	✓	✗	✗	✗	✓
Identify child is in proximity to hazardous object	✗	✗	✗	✗	✓
Prevent accidents before happening	✗	✓	✗	✗	✓

Comparison of Similar Research

This research aims to develop a system to identify possible hazardous objects effectively and accurately in a child area and take responsive actions to prevent or minimize the danger by utilizing computer vision and deep learning approaches with a combined video surveillance system.

1.3. Research Problem

For parents, their children are the most important thing in their lives. It is becoming harder for parents to take proper care of their children 24 hours a day. Children often need continuous elderly attention, or otherwise, there is always a risk of them getting unintentionally injured, as they are not psychologically developed enough to identify dangers at an early stage.

The statistics obtained from worldwide studies on children reveal an astonishingly high number of annual child injuries and subsequent deaths. The organization 'Safe Kids Worldwide' indicates that more than 2,200 children die annually from domestic injuries. There can be many reasons for this to happen. Some can be directly related to parents or such authorities such as,

- Lack of parents' attention towards children
- Parents not having time to spend with children because of work
- Leaving children with outsiders
- Children being very disobedient due to the absence of a strong relationship between parents and children.

As much as parents like to spend time with their children, it might not always be possible, especially when both have to work to earn a living for the family. Thus, there are unfortunate times that children face dangers when parents are not around. Sometimes the reason could be the environment in which the child is placed. Some could be stated as,

- Unsafe play areas
- Unsafe housings and household furniture
- Lack of proper lighting
- Lack of proper wastage management
- Careless handling of harmful objects and materials
- Careless usage of kitchen utensils and such equipment

Even if parents or some elderly parties were around the child, it is hard to keep tailing a child. A toddler that advances into walking from crawling needs to be given their own space for them to explore the environment and learn. This is crucial for their physical and psychological development. Thus, it becomes even harder to protect a child from sudden injuries during this period of development.

WHO statistics show that in Europe alone, 26,000 children die from injuries every year. The rate is equivalent to 3 child deaths per hour. That study further shows that out of the injuries reported in a year, around 94% is reported as cuts, falls, burns and collisions.

This raises a significant concern regarding the safety of the environment that the children live in. Cuts, burns, and collisions that children encounter are mostly caused by sharp, hot or heavy objects commonly found around a household. Such objects might be scissors, knives, hot water kettles, hot cups, heavy televisions and other heavy furniture. Objects of these types can be recognized as 'hazardous objects' to a child.

Even if parents pay close attention to their children, it might not always be practical to 'childproof' the whole environment. There are many child-proofing practices currently adopted such as installing safety gates, baby bumpers and baby room cameras, but the effectiveness of these methods are highly questionable as WHO reports the child death rate as hardly changed. In a normal household with working parents, anybody can accidently leave any hazardous objects in a rush, where children can have access to it. When a toddler finds any object, it is common for them to be curious about the object and explore it. But they would not know how it works and how it can harm them until they get injured by themselves.

Death from an injury is not the only concern. Even if the death was prevented, severely injured children might be left physically and/or mentally damaged for the rest of their lives. Mind-shattering accidents that happen at the early stage of the learning and development process might negatively affect the child's future development. Long-term physical disabilities might occur, preventing future physical evolution. Long-term psychological disorders might cause a child to be afraid of physical activities even if their physical status is in good condition.

Safety gates, bumpers or monitors do not provide parents or authorities with real-time danger alerts if the child somehow encounters one. These methods, as it seems, provide only danger avoidance, not effective prevention. Hence, it is very important and timely to implement a more effective system to prevent hazardous events from happening and protect children from hazardous objects.

2. OBJECTIVES

2.1. Main Objectives

The main objective of the proposed solution is to help to develop a safe environment for children to ensure safety at the most important time period of growth in their life while addressing the key failures of existing child protection mechanisms.
The solution aims to establish a safer environment, minimizing the danger of children getting injured by hazardous objects during the critical stage of child development.

The solution will be implemented as a system that detects hazardous objects within reach of a child, using surveillance cameras and taking appropriate actions to alert a responsible party.

2.2. Specific Objectives

To realize the main objective of the proposed solution, the following specific objectives must be realized in parallel.

- Detecting hazardous objects and children using a live video feed.

  The system should be capable of processing live video feed to detect objects and children within the area. It should specifically be able to detect children, not adults, in order to be the most effective. It should not be processing video feed when a child is not present in the area, as it would be a wastage of expensive computer resources.

- Identifying if the child is in proximity to a hazardous object.

  When both the child and the hazardous object is identified, the system should be able to identify the distance between them.

- Identifying the danger level of the hazardous object.

  The system should be capable of identifying if the child is within a dangerous distance form the object. The system should continuously compare the actual distance with a pre-determined distance fed to the system to accurately identify whether the distance is dangerous or not. Also, according to that pre-determined distance level, the system should identify the danger as low-risk or high-risk.

- Trigger the appropriate warnings.

  Finally, as the responsive action, the system should trigger appropriate warning alerts. If the distance from the child to the object is not dangerously close, the system should trigger an alarm aimed at the child, to divert his attention from the object. If the distance is close, as per the determination of parents, the system should alert them or any other responsible party about the danger.

The proposed system would not aim to prevent hazardous event from happening altogether, as it depends on the response from the parents after receiving the alert. However, it would be more effective than existing mechanisms in avoiding dangers with actively alerting responsible elders.

3. METHODOLOGY

During the implementation process of the systems, the key steps we followed are

1. Requirement Gathering
2. Data collection
3. Preprocessing and annotation
4. Training detection model
5. Testing the model

3.1. Requirement Gathering

The research domain and background were analyzed in requirement gathering, leading to a complete study on child death and injury records of past years, possible causes, existing mechanisms to prevent child injuries, and limitations.

After analyzing the requirement, the current processes established as solutions to the domain problem and other similarly implemented systems were studied to understand existing solutions. Defining the project scope and understanding the domain I would be covering was a challenge, as the scope covers a broad aspect of child care. Regarding the proposed application, interviews were conducted with the parents and caregivers of children to find out the most appropriate features to implement for this.

Key procedures that were followed when obtaining reliable requirements and understanding the project scope for the problem domain can be stated as follows.

- Collecting related research papers
- Feasibility study
- Background and literature survey
- Reading and analyzing collected research papers
- Examining financial viability.
- Collecting data from users and analyzing their perspectives about this system
- Filtering out the most relevant components and finalizing the project scope.

3.2. Data Collection

Building a computer vision model from scratch would be a highly time consuming, excessive amount of work. Image classification is challenging because of the problems like view-point variation, scale variation, illumination and background clutter, which would not bother human vision but would bother computer vision when identifying objects.

To overcome this challenging situation, we adopted the transfer learning method to develop the system. Transfer learning allows us to pre-train a computer vision model using training image data sets and use that trained model to solve the actual problem.

Training the data set is the most essential and challenging part of developing a custom computer vision model. Since the training network is a deep neural network, it takes an enormous amount of image datasets to perform specific tasks. To develop the detector, the dataset was obtained by extracting the frames consisting of hazardous objects which were collected from multiple sources. The footage from home surveillance cameras which were available on the internet, CCTV camera footage captured by ourselves, YouTube videos and Google images contributed to the image obtaining process resulting in a total of 3234 images.

The collected images were classified into four image classes: Tea Cup, Scissors, Knife and Gas Kettle.

fig 2 1. Image Classes (Knife, Scissor, Teacup, Gas Kettle)

During the performance testing phase of the model, recorded videos and real-time video feeds were used. The detection was carried out in the form of a test and classified the false positives and negatives of the detector.

3.3. Data preprocessing and annotation

The Initial step was to annotate the obtained data in order to perform transfer learning on the model. Image annotation is the process of labelling images of a dataset in order to train a machine learning model. Image annotation is used to label the hazardous objects before training the detection algorithm. As a preprocessing step, the gathered different resolution images were converted into a single frame size for the ease of training the model. The annotation phase consumed five days for all the collected images, as we had to do the annotation on the whole training dataset manually.

This annotation was the bounding box of the hazardous objects available in the images. The bounding box marks the height, width, and depth of an object and helps derive the object's coordinates in the space. The annotation was done with the assistance of the "Lableimg" tool. Labeimg is a widely-used, open-source graphical image annotation tool written in Python and uses Qt for the graphical interface. This tool is very useful for object localization and detection tasks.

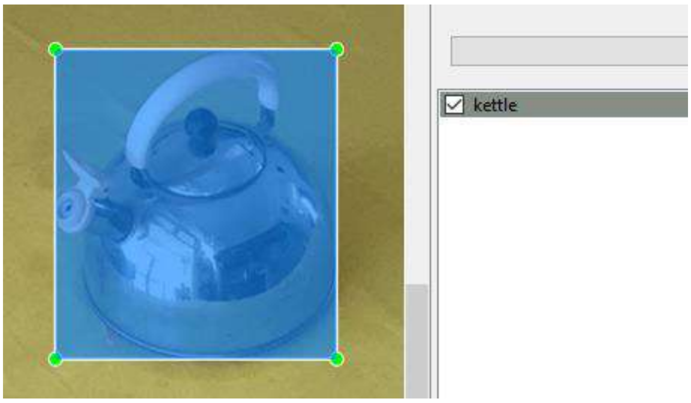

fig 2 2. Annotation

LabelImg is only able to create annotation boxes around considered objects. For that specific task, LabelImg has all the necessary functionalities, simplifying the task.

Furthermore, LabelImg provides the final output in 3 popular annotation formats, PASCAL VOC, YOLO and CreateML. The result consists of the x, y coordinates of the centre of the bounding box and the width and height of the bounding box.

As a demonstration of the result of the annotation, the following can be obtained.

- Size of the bounding box, including width, height and depth

```
<size>
        <width>1200</width>
        <height>1200</height>
        <depth>3</depth>
</size>
```

fig 2 3. Size of bounding box

- coordinates of the object as for its placement in the space

```
<bndbox>
        <xmin>496</xmin>
        <ymin>658</ymin>
        <xmax>523</xmax>
        <ymax>778</ymax>
</bndbox>
```

fig 2 4. Bounding Box Coordinates

3.4. Model Training Environment

The model training was carried out on the Google Colab environment. Google colab is a free and open-source product from Google. Colab is a web-based tool that allows anyone to write and run Python code, and it's especially useful for machine learning and data analysis. Colab is a hosted Jupiter notebook service that doesn't require installation and gives you free-of-charge access to computing resources like GPUs. The training has been broken down into a few steps as following.

- Configuration of the GPU environment.
- Install Darknet YOLOv4 training environment.
- Download the custom dataset which mentioned in the 'Dataset' section for YOLOv4 and set up the directories.
- Configure a custom YOLOv4 training config file for Darknet.
- Train our custom YOLOv4 object detector.
- Reload YOLOv4 trained weights and make inference on test images

3.5. Detection Model Selection for Hazardous Object Detector

YoloV4 and YoloV4 tiny are the two data models based on the darknet framework, trained in the process of developing the hazardous object detector. Darknet is an open-source neural network framework written in C and CUDA. It is fast, easy to install and supports CPU and GPU computation. Darknet can be installed with an OpenCV dependency which supports a wider variety of image types. Darknet can be used to classify images as it features YOLO (You Only Look Once), a state-of-the-art, real-time object detection system. However, the XML/VOC files resulting in the annotation process using the LabelImg tool does not support the darknet framework of the YOLO data model. Hence, To convert annotated image files into YOLO Darknet supported annotation, Roboflow software is used.

The two detection models used for the case are Yolov4 and Yolov4 Tiny, based on the Darknet Framework implementations. The purpose of using these two models is to select the most accurate and efficient model for developing the detector.

These models are pre-trained on the MS-COCO dataset which is used for various computer vision projects, whereas the detector generally learns to detect objects. Machine Learning and Computer Vision engineers popularly use the COCO dataset for various computer vision projects. COCO (Common Objects in Context) is large-scale object detection, segmentation, and captioning dataset published by Microsoft. COCO provides several helpful features such as:

- Object segmentation
- Recognition in context
- Superpixel stuff segmentation
- 330K images of which over 200k is labelled
- 1.5 million object instances
- 80 object categories
- 91 stuff categories
- Five captions per image

As the image dataset was created with the goal of advancing image recognition, the COCO dataset contains challenging, high-quality visual datasets for computer vision, primarily state-of-the-art neural networks. For example, COCO is often used to benchmark algorithms to compare the performance of real-time object detection. The format of the COCO dataset is automatically interpreted by advanced neural network libraries. In computer vision, transfer learning is usually expressed through the use of pre-trained models. The trained model is working as a detector as the result of the transfer learning solution. In Computer vision, transfer learning is a popular method since it is allowed to build more accurate models.

There are two approaches have been performed using detection models

1. Identify the possible hazardous objects (Knife, Scssiors, Gas Kettle, Teacup)
2. Adault child classification

This report is mainly represented the hazardous object detection methodology. To perform efficient and accurate model, ReLu (Rectified Linear Unit) was used and adam was used as a model optimizer. To get a better accuracy the following methods were used.

1. Changed the number of epochs: The number of epochs is a hyperparameter that controls how many times the learning algorithm runs through the full training dataset.

2. Changed the batch size: The batch size is a hyperparameter that specifies how many samples must be processed before the internal model parameters are updated.

When considering the output of our model, we focus on optimizing it for sum-squared error. We utilize sum-squared error due to the fact that it is simple to improve. However, it falls short of our goal of optimizing average precision. It gives equal weight to localization and classification errors, which is not ideal. In addition, many grid cells in each image do not include any objects. This reduces the' "confidence" scores of those cells to zero, often overthrowing the gradient from cells that contain objects. This can cause model instability, leading early training divergence.

3.6. Detection Model Selection for Child Adult Classification

We used image processing techniques to build a set of processes to accomplish real-time adult-kid classification. When training and deploying real-time adult child categorization models, the Kaggle dataset was used. The total number of photographs in the collection was 4754. The dataset was divided into two categories: child and non-child. The chosen dataset included a mixture of kid and non-child photos, totaling 2168 and 2586 images, respectively. The photographs collected included a variety of cultural child images. It has had an effect on the model's accuracy.

Keras API was utilized to create the dector model. This is due to the fact that Keras being a simple framework for quickly defining deep learning models based on TensorFlow. Keras is a fantastic platform for deep learning applications too.

The training and testing data folders were initially supplied to the keras library generators with the data labeled as child and non-child. Following labeling, these generators were used to generate models for VGG16, InceptionV3, and ResNet50, among other transfer learning models. An accurate model was created using the Rectified Linear Unit (ReLU) technique. The role of optimizer was assigned to Adam. The number of epochs and training/testing batch sizes were modified to improve accuracy. We have also added some dense layers to each of the three models. Two completely connected layers and one output layer make up the added dense layers. The output layer has a sigmoid function, and all have RELU activation. The first dense layer consists of 1024 units, the second layer consists of 512 units, and the last layer has only one unit for being the output layer. TABLE 1 shows the results that were achieved.

TABLE 1. Model Accuracy of Child Adult Classification

Model	No of Epoch	Batch size	Added Layers	Accuracy
VGG16	25	32	Three dense	52.65%
ResNet50	25	32	Three dense	92.99%
Inceptionv3	25	32	Three dense	94.42%

3.7. How the detector works

The detector mainly aims to provide a solution for the fact that children being unable to identify the danger associated with a hazardous object and the fact that parents or elders being unable to always stay with the child giving them protection against such hazardous objects. The dangerous results occurred for the kids due to the use of hazardous objects and not having the capability of distinguishing the hazardous objects while measuring the danger level of the objects takes a major count.

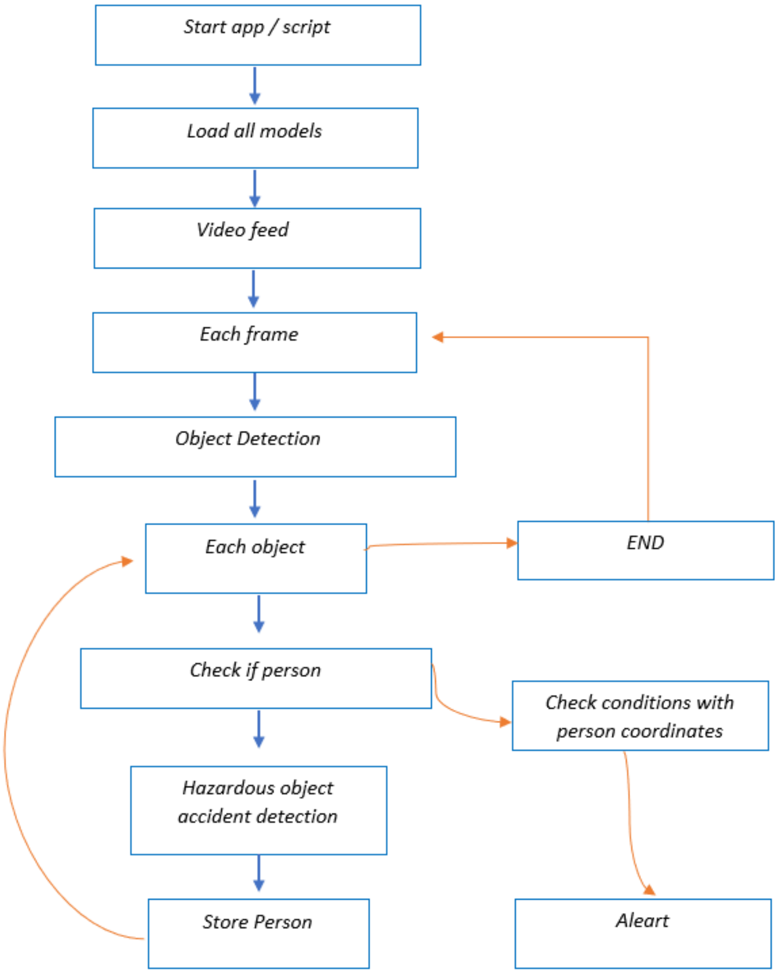

fig 3 1. High level diagram

According to above high level system architecture, system input is a real time video feed and output is a alert whether the child is in proximity to the hazardous object or not.

This research is mainly based on object detection, which is also a combination of three functions; Object recognition, to find objects in an image, Object localization, to find where exactly in the image the objects are located; and Object classification, to detect what particular objects are in that image.

fig 3 2. Image Classification vs. Object Detection tasks

This paper is based on the YOLO detector model, which can predict the object class and the probability of object class in the bounding box. Below mentioned parameters have each bounding box.

1. The centre position (dx, dy)
2. The height of the bounding box (dh)
3. The weight of the bounding box (dw)
4. The object class (c)
5. A probability value (pc)

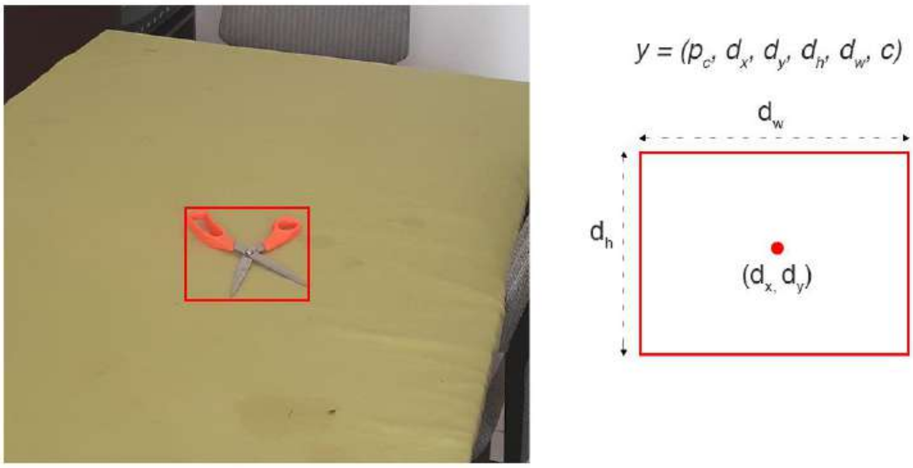

fig 3 3. Bounding Box Coordinates

A probability value (Pc) is the probability of an object class in that bounding box. An input image is split into several grid cells, typically using a 19x19 grid. Each cell is responsible for predicting five bounding boxes, and there can be one or more objects in a cell.

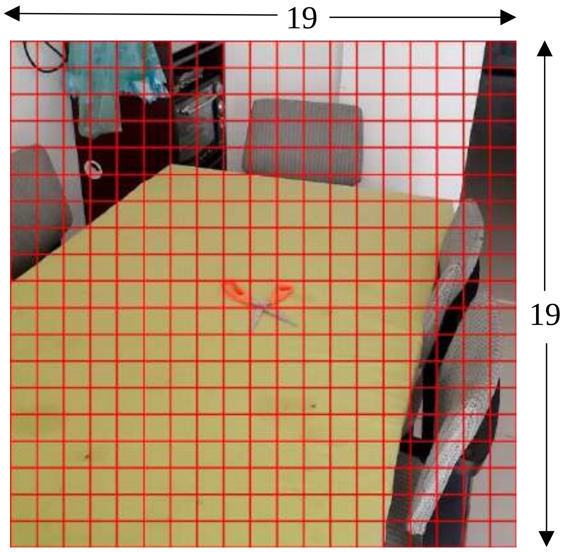

fig 3 4. 19x19 Grid view

Most of the bounding boxes in the cell may not have an object. The filtration of these bounding boxes is done based on object class probability value (Pc). When performing the object detection algorithm (YOLOv4 and YOLO Tiny), it is predicted that different size several bounding boxes for the same object. Fig

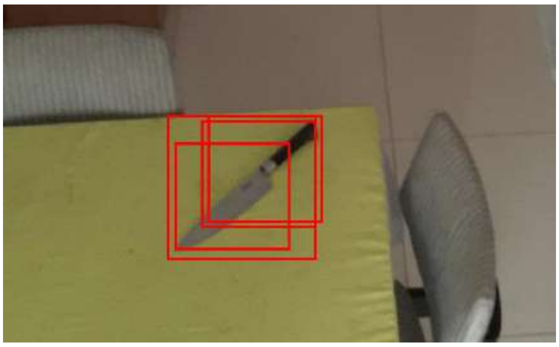

fig 3 5. Before non-max suppression

To eliminate theunwanted bounding boxes, the non-max suppression is used and only the highest probability bounding box will remain.

fig 3 6. After non-max suppression

The detector is trained with 4 types of image classes which are Teacup, Scissors, Knife and Gas Kettle. These image classes are created as an initial step to recognize 'sharp' and 'hot' objects.

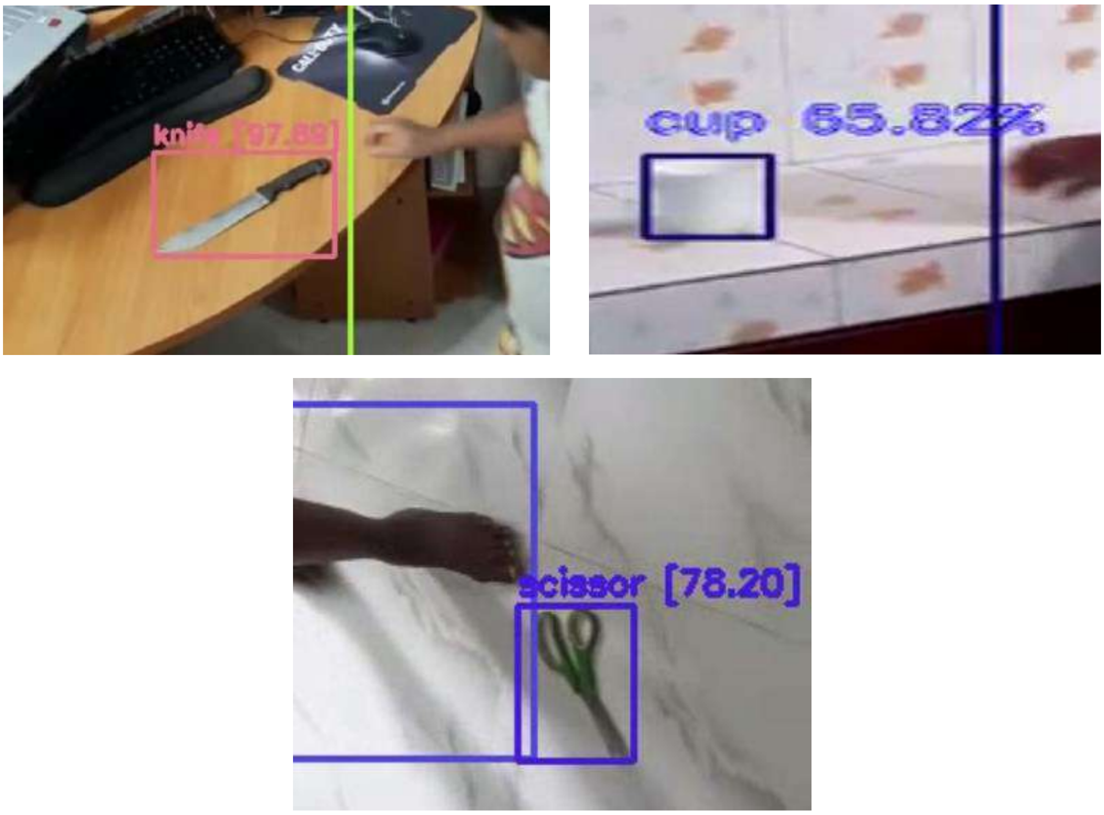

fig 3 7. Image Classes Trained using YOLOv4

3.8. Prerequisites

Bounding box Overlapping

This method is used to identify whether the child is in proximity to the hazardous object or not. According to the result outcomes of this method, the relevant alerts will trigger. To perform this method, child and object bounding boxes are used. The maximum distance (Safe distance) between the child and the hazardous object is pre-defined by the user. The system always continues to check the distance between the two bounding boxes, compared against the pre-defined distance. If the child comes towards the hazardous object, system identifies it as a breach of the safe distance; an overlapping of the bounding boxes.

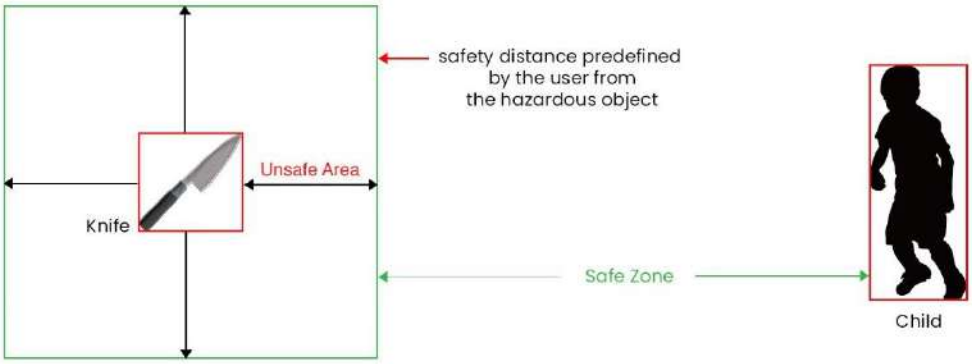

fig 3 8. Child is in safe area, as he is out of the pre-defined distance

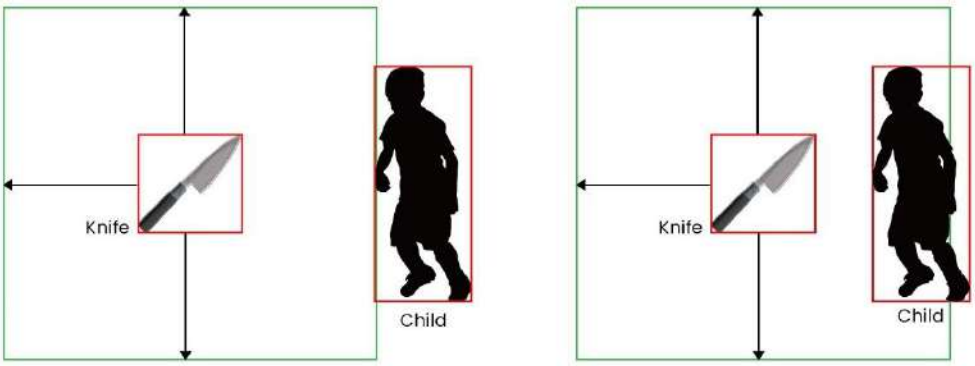

fig 3 9. Child is in the danger area, as the child bounding box either touches or overlaps the object bounding box

This logic is implemented by using a calculation to measure the width and the 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

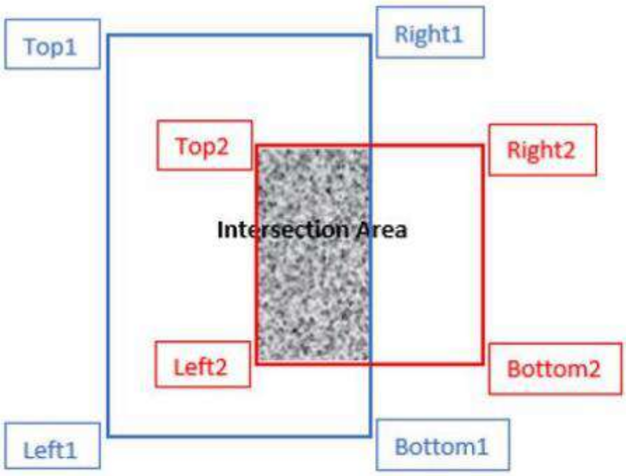

fig 3 10. Intersection area represent as a danger area

Intersection is represented by the overlapped area. Following calculation is performed for calculating the overlap area

$x = max (Left1 , Left2)$
$y = max (Top1 , Top2)$
$width = min (Left1 + right1 , Left2 + Right2) – x$
$height = min (Top1 + Bottom1 , Top2 + Bottom2) – y$

```
# Calculate the intersection between two bounding boxes
def intersection(a,b):
  x = max(a[0], b[0])
  y = max(a[1], b[1])
  w = min(a[0]+a[2], b[0]+b[2]) - x
  h = min(a[1]+a[3], b[1]+b[3]) - y
  if w<0 or h<0: return False # or (0,0,0,0) ?
  return True
```

fig 3 11. Implemented calculation to find the overlap area

According to the intersection value using above mentioned calculation, the following alerting methods have been implemented.

1. No alerting when the child is out of close proximity to the object, decided against the pre-defined distance of danger.
2. Distraction alert to the child when the child begins to cross the predefined distance.
3. Different distraction alert to the child when the child is within 75% of the danger distance to the object.
4. Danger alert to the parent when the child is within 50% of the danger distance.

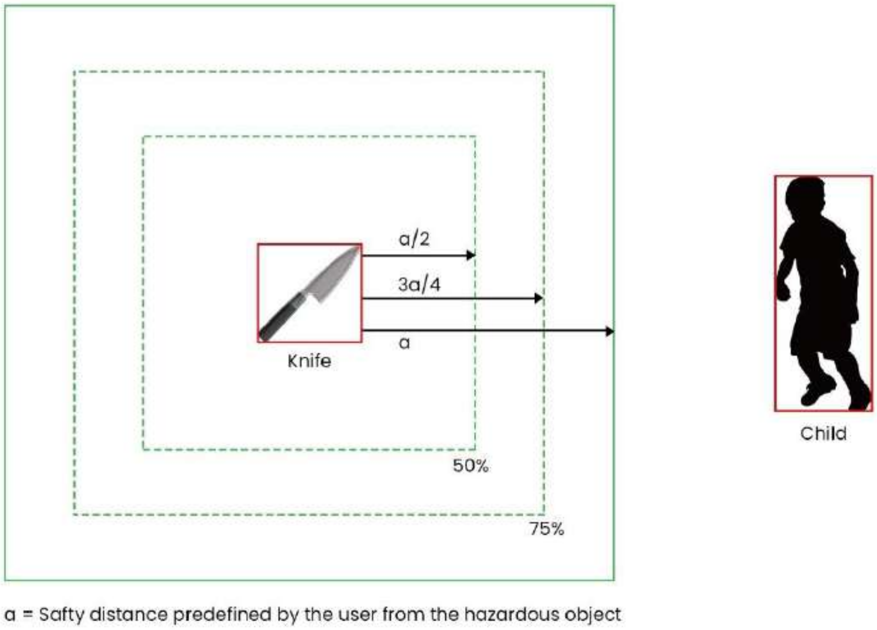

fig 3 12. No alerts will trigger. Child is in safe zone

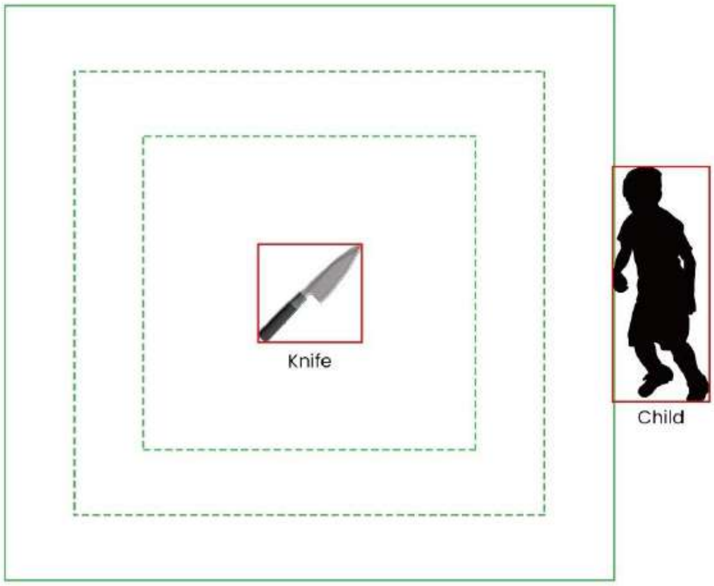

fig 3 13

An alert towards the child will be triggered when the child bounding box touches the object bounding box as the child enters the danger zone.

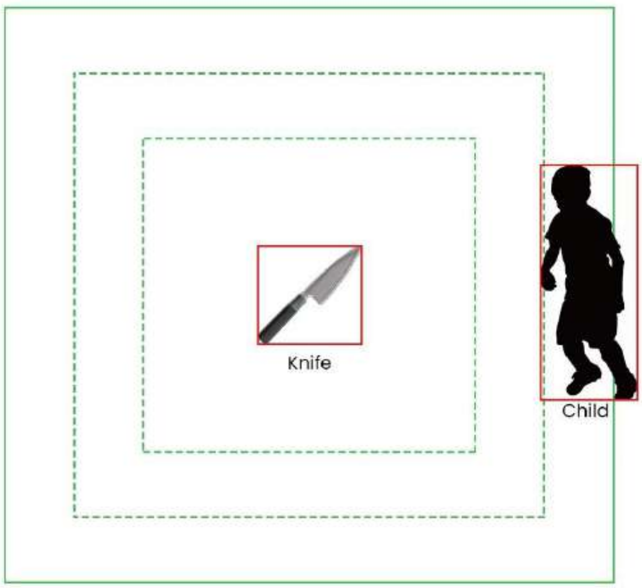

fig 3 14

Different Child Distract alert will trigger. Danger medium

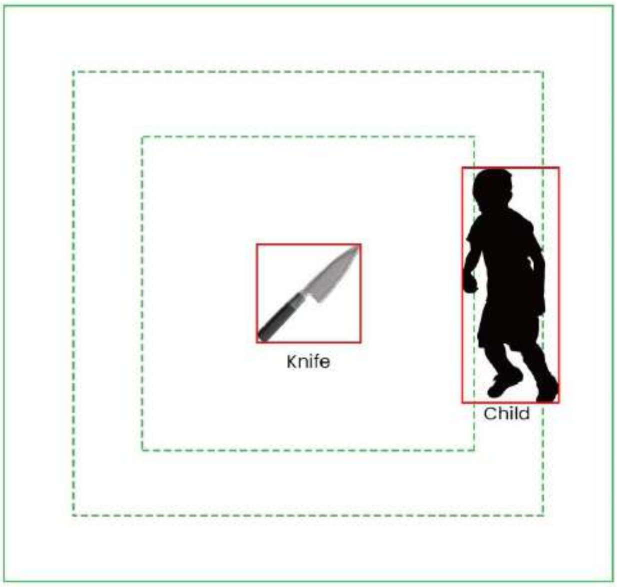

fig 3 15

An alert to the parent is triggered when the child is within 50% of the danger distance to the object, identified as high danger level.

3.9. System Implementation Steps

Initially, the detector detects hazardous objects and identifies them according to the above-mentioned image classes. In parallel with that process, it detects where the hazardous objects are located and marks the bounding boxes, and does the same for the child.

The danger distance which is carried out as the major detecting fact is a predefined variable by the user. According to that danger distance, the detector defines the distance between the child and the object and the following actions for which the trigger warnings will be executed.

When a child is entering the frame, the detector identifies the child as the other object ignoring the parents and any caregivers affecting the scenario and marks the bounding box of the child. It should be noted that the detector would process the video feed only when there is a child detected within the area, or otherwise the computer resources used to continuously process the video feed would be in vain.

The YOLOv4 model is trained to identify the objects and it was selected due to the fact of processing the real-time data and producing the most accurate results as the outcome.

Once both the bounding boxes of the child and the hazardous object are clearly detected, the detector works in real-time and checks whether the two bounding boxes are in a proximity level. The proximity level is tracked by the detector with the real-time process of detection by repeatedly comparing with the predefined danger level, in order to execute the trigger warnings of the detector.

The trigger warning is executed along with the proximity level of two bounding boxes. With the purpose of distracting the child from the hazardous object, the triggering will be executed if the proximity level is low. The high proximity level will execute the trigger warning which is going to inform the parent or caregiver about the danger.

Here, the 'low-level' of danger and the 'high-level' of danger would also be defined by the system according to the pre-specifications made by relevant authorities.

4. RESULTS AND DISCUSSION

The above-mentioned test data was used to evaluate the detector's performance. This section mainly focuses on the results and discussion of the hazardous object detection and accident prevention component. This section provides a summary of model training results and figures and outcomes gathered throughout the testing phase

4.1. Model Training Results

Hazardous object detection models have been implemented by training the YOLOv4 and YOLOv4 Tiny models, in order to compare the performance of both the models and choose the best one. Training data were fed to those two models as 75% and 25%, respectively, and the testing data also has been divided accordingly. The result comparison of the two model is reflected in table

TABLE 2. Model training results of Hazardous objects

Model	Avg.Precision	True Positive	False Positive	No. Epochs
YOLOv4	94.06%	272	78	1000
YOLOv4 Tiny	91.20%	248	29	8000

According to the results obtained while testing the previously trained models, YOLOv4 has seemingly performed attributing to high accuracy and fast responsiveness for hazardous object detection. YOLOv4 has also responded well for the low lighting conditions.

The following are the results for the YOLOv4 model built for the hazardous object (Knife, Scissors, Gas kettle, Teacup)

TABLE 3. YOLOv4 Model training resuls for particular object class

Class Name	Avg. Precision
Knife	97.27%
Kettle	87.21%
Scissor	96.30%
Teacup	95.48%

YOLOv4

The following are the results for the YOLOv4 Tiny model built for the hazardous object (Knife, Scissors, Gas kettle, Teacup)

TABLE 4. YOLOv4 Tiny Model training resuls for particular object class

Class Name	Avg. Precision
Knife	99.70%
Kettle	98.23%
Scissor	85.90%
Teacup	80.98%

4.2. Hazardous Object Detection Model Performance Evaluation

During the performance testing phase of the model, 20 recorded video clips and realtime video feeds were used which consist of different angles, low light conditions and law quality videos, to ensure that the detection approach outputs acceptable results as expected, in many conditions as possible. The results are evaluated according to the following categories.

a. True Positive (TP) = Child is closer to the hazardous object and system detected as a danger

fig 3 16. Case 1 - True Positive (TP)

b. True Negative (TN) = Child is not closer to the hazardous object and system detected it as not a danger

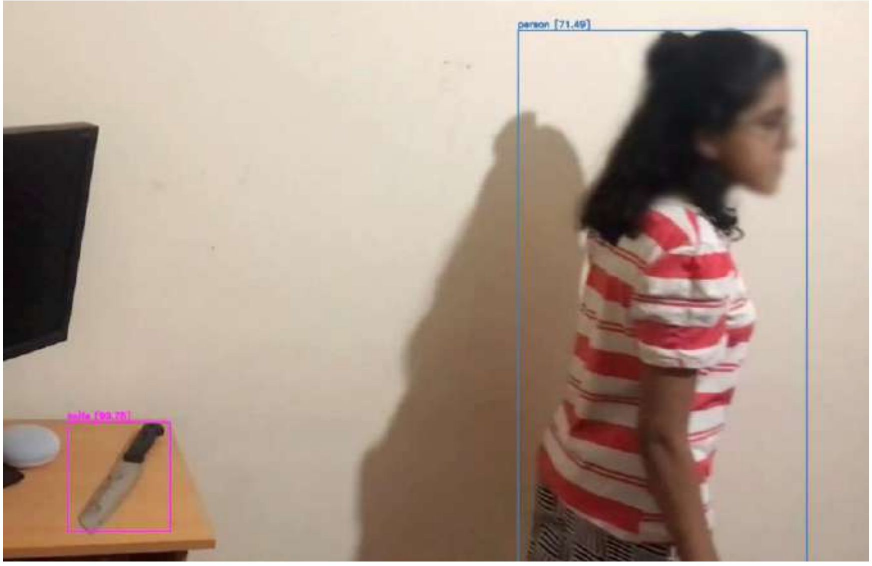

fig 3 17. Case 2 - True Negative (TN)

c. False Positive (FP) = Child was going away from the hazardous object, and system detected as a danger

fig 3 18. Case 3 - False Positive (FP)

d. False Negative (FN) = Child was closer to the hazardous object, and system detected as not a danger

fig 3 19. Case 4 - False Negative (FN)

TABLE 5. Performance evaluation table

NO.Recored Video = 20	Status : Danger	Status: Not danger
Detected: Danger	10 (TP)	2 (FP)
Detected: Not danger	3 (FN)	5 (TN)

We used a set of criteria to assess this approach, including accuracy, sensitivity, and specificity [32]. The equations below were utilized.

Sensitivity = TP/(TP + FN) = (Number of true positive assessment)/(Number of all positive assessment)

Specificity = TN/(TN + FP) = (Number of true negative assessment)/(Number of all negative assessment)

Accuracy = (TN + TP)/(TN+TP+FN+FP) = (Number of correct assessments)/Number of all assessments)

fig 3 20. Accuracy, sensitivity, and specificity equations

Using the above equations, the actual hazardous object detection and accident prevention component performance is calculated

Sensitivity (%) = TP/(TP + FN) x 100 = 10/(10 + 3) x 100 = 76.92%
Specificity (%) = TN/(TN + FP) x 100 = 5/(5 + 2) x 100 = 71.42%
Accuracy (%) = (TN + TP)/(TN + FN + FP + TP) x 100 = 15/20 100 = 75%

4.3. Results Summery

Below figures have obtained the testing phase of hazardous object detection and accident prevention component

As these figures clearly show, when a child is entering the frame, the detector identifies the child as the other object, ignoring the parents and any caregivers affecting the scenario, and marks the bounding box of the child. It should be noted that the detector would process the video feed only when a child is detected within the area, or otherwise the computer resources used to continuously process the video feed would be wasted pointlessly.

fig 3 21. Only object is in the video feed

Fig 3.18 and fig 3.19 show an instance where both the child and the hazardous object are detected while the both appear in one frame.

fig 3 22

fig 3 23

Check the proximity and trigger alerts accordingly

Fig 3.20 that the child is in the safe zone, decided against the provided safeguard distance. Thus, an alerting would not be triggered in this situation.

fig 3 24

Fig 3.21 shows, a child coming towards the hazardous object. In this case, system triggers an alert aimed at the child, to distract the child away from the object in focus. Danger level is medium in this case. Bothe the child and the hazardous object (scissor) were identified successfully. Main goal of this step is to prevent the accident by distracting the child at a safer distance.

fig 3 25

Fig 3.22 and fig 3.23 show, a child being closer to the hazardous object. Here, the system triggers an alert to the parent to inform them about the danger. Danger level is high in this case. Both the child and the hazardous object (scissor) were identified successfully. This step is taken, so as to minimize the danger to the child by notifying parents

fig 3 26

fig 3 27. Hazardous object detector support different type knife

Hazardous object detector is trained to recognize different items of the same category. The figure above shows the detector supporting a different type of a knife.

Fig 2.24 shows, the detector supporting low lighting conditions. Even in this case, both the child and the object detection were successfully completed along with proximity alerts

fig 3 28

Figuar shows, the hazardous object is difficult to be detected when it is covered by the child.

fig 3 29

4.4. System Requirement

Functional Requirements

- Should have a method to localize the hazardous objects in child area.
- Should have a method to classify the hazardous object danger level (Hot liquid container has high risk)
- Should have a method to detect Child's actions without identifying parents' actions.
- Should have a method to classify whether Child is in proximity to a hazardous object.
- Should have a method to distract child activity or alert to the parents or caregiver about the danger.

Non-Functional Requirements

- System efficiency and performance
- Availability
- Accuracy

Personal Requirements

- Parent/Guardian should be available.
- A 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.

5. COMMERCIALIZATION

In a society where families with both parents working has become a common norm, c hildren 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.

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.

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.

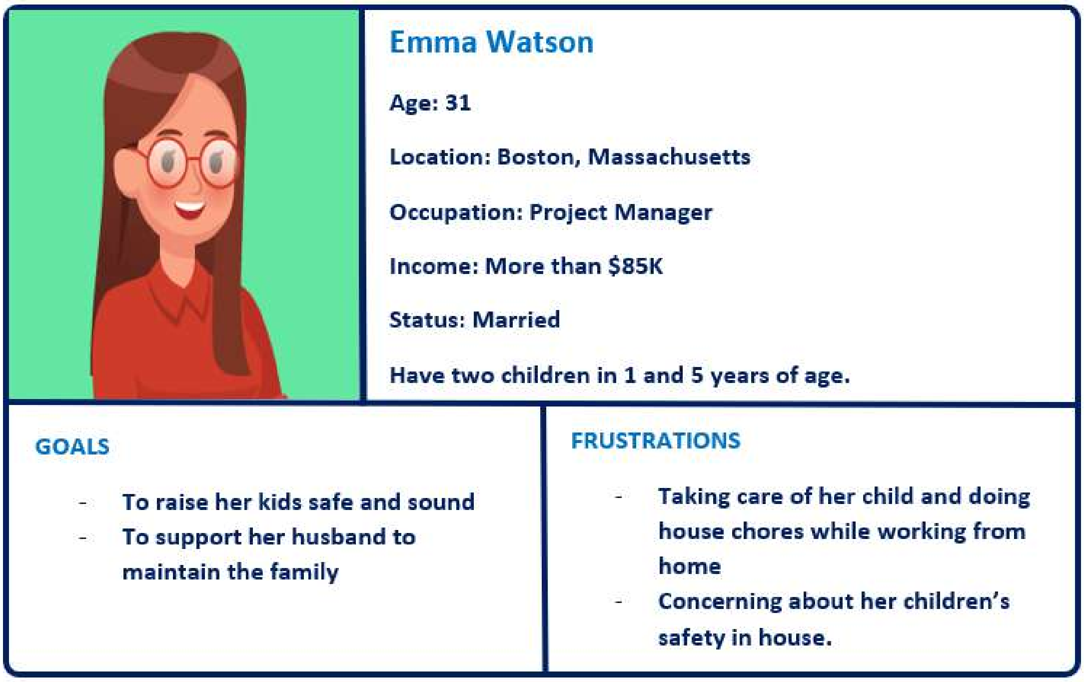

Fig. 7. User persona of working from home parent

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.

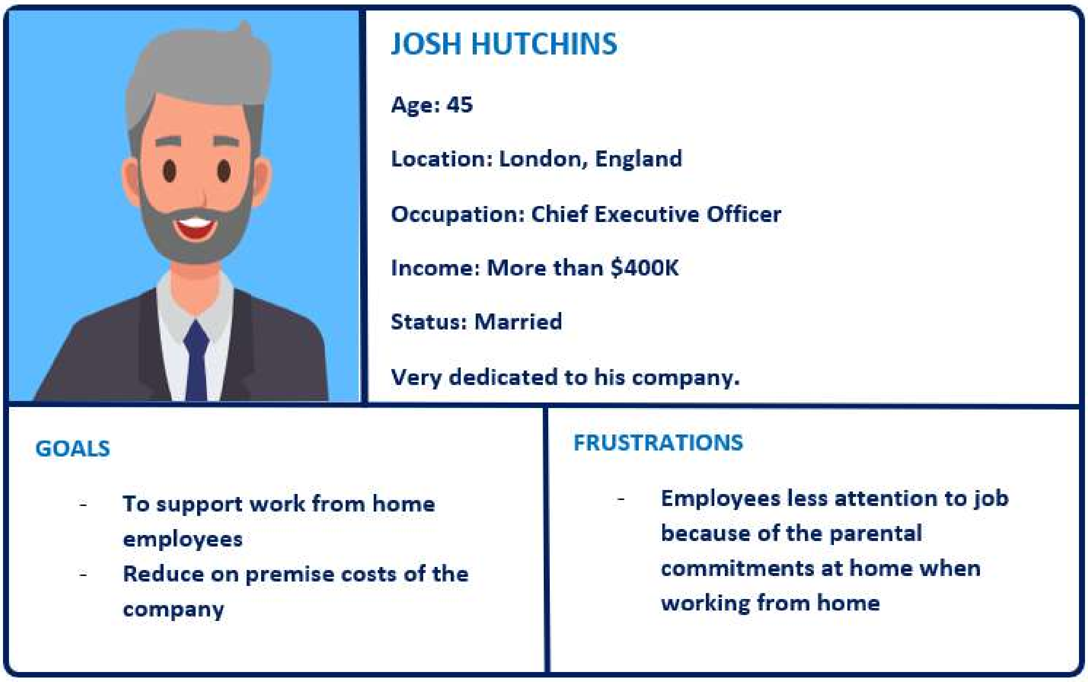

Fig. 8. User persona of a chief executive officer of a company supporting working from home

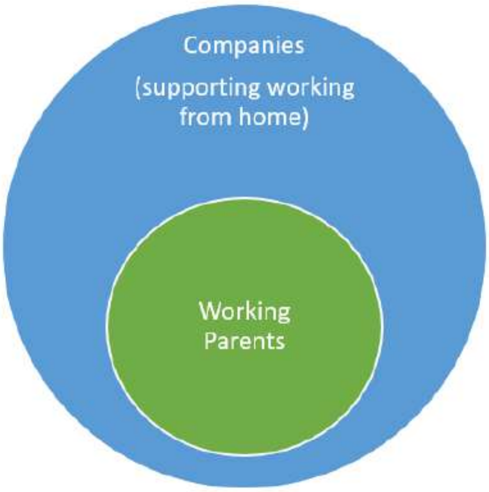

Fig. 9. Target market for AICare

5.2.1. 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.

	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.

Fig. 10. Competitor analysis

5.2.2. 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.

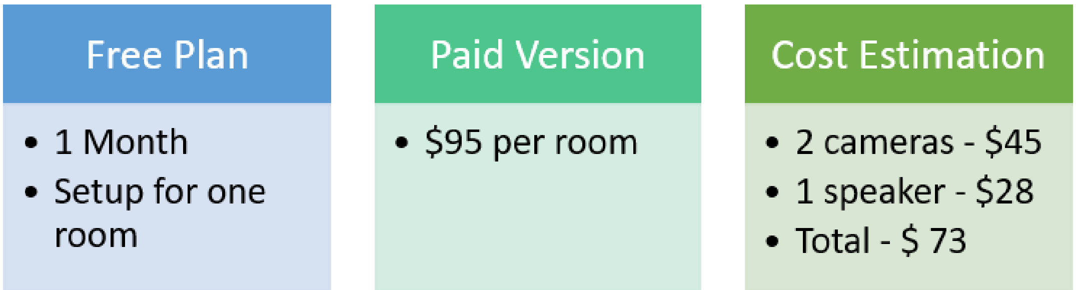

Fig. 11. Business model of AICare

6. TEST CASES

Test Scenario ID	Sharp Object -1		**Test Case ID**	Sharp object -1A		
Test Case Description	Sharp Object – Positive test case		**Test Priority**	High		
Pre-Requisite	NA		**Post-Requisite**	NA		
Test Execution Steps:						
S. No	**Action**	**Inputs**	**Expected Output**	**Actual Output**	**Test Result**	**Test Comments**
1	Show Knife at the Webcam	Image frames of the Knife	Detect and Highlight the Knife on the screen	Detect and Highlight the Knife on the screen	Pass	Detect Sharp object successful

Test Scenario ID	Sharp Object -1		**Test Case ID**	Sharp Object -1B		
Test Case Description	Sharp Object – Negative test case		**Test Priority**	High		
Pre-Requisite	NA		**Post-Requisite**	NA		
Test Execution Steps:						
S. No	**Action**	**Inputs**	**Expected Output**	**Actual Output**	**Test Result**	**Test Comments**
1	Show Iron Rod at the Webcam	Image frames of the Iron Rod	Does not Detect or Highlight the Iron Rod on the screen	Does not Detect or Highlight the Iron Rod on the screen	Pass	Does not Detect non-Sharp object successful

Table 7.2 – Test Case 2

Test Scenario ID	Hot Liquid Container -1	**Test Case ID**	Hot Liquid Container -1A
Test Case Description	Hot Liquid Container – Positive test case	**Test Priority**	High
Pre-Requisite	NA	**Post-Requisite**	NA

Test Execution Steps:

S. No	**Action**	**Inputs**	**Expected Output**	**Actual Output**	**Test Result**	**Test Comments**
1	Show a Kettle at the Webcam	Image frames of the Kettle	Detect and Highlight the Kettle on the screen	Detect and Highlight the Kettle on the screen	Pass	Detect Hot Liquid Containers successful

Table 7.3 – Test Case 3

Test Scenario ID	Hot Liquid Container -1	**Test Case ID**	Hot Liquid Container -1B
Test Case Description	Hot Liquid Container – Negative test case	**Test Priority**	High
Pre-Requisite	NA	**Post-Requisite**	NA

Test Execution Steps:

S. No	**Action**	**Inputs**	**Expected Output**	**Actual Output**	**Test Result**	**Test Comments**
1	Show water bottle at the Webcam	Image frames of the Water Bottle	Does not Detect or Highlight the Water Bottle on the screen	Does not Detect or Highlight the Water Bottle on the screen	Pass	Does not Detect non-Hot Liquid Container successfully

Table 7.4 – Test Case 4

Test Scenario ID	Sound Alert -1	**Test Case ID**	Sound Alert -1A
Test Case Description	Sound Alert – Positive test case	**Test Priority**	High
Pre-Requisite	NA	**Post-Requisite**	NA
Test Execution Steps:			

S. No	**Action**	**Inputs**	**Expected Output**	**Actual Output**	**Test Result**	**Test Comments**
1	Video of a child coming to close distance for a knife	Image frames of the Video	Play warning recording to the child in the speaker	Play warning recording to the child in the speaker	Pass	Give sound alert successful

Table 7.5 – Test Case 5

Test Scenario ID	Sound Alert -2	**Test Case ID**	Sound Alert -2A
Test Case Description	Sound Alert – Positive test case	**Test Priority**	High
Pre-Requisite	NA	**Post-Requisite**	NA
Test Execution Steps:			

S. No	**Action**	**Inputs**	**Expected Output**	**Actual Output**	**Test Result**	**Test Comments**
1	Video of a child coming to close distance for a kettle	Image frames of the Video	Wail Siren	Wail Siren	Pass	Give sound alert successful

Table 7.6 – Test Case 6

7. CONCLUTION

The principal focus of the research was to detect, prevent and minimize the danger of unintended child accidents, caused by improper handling of household hazardous objects by children. The proposed solution was implemented by incorporating home surveillance cameras along with computer vision models and an alerting system to acquire the desired outcome. The danger was identified by first detecting the child and possible hazardous objects in an area and secondly detecting the distance between the child and the object. The model is capable of detecting four types of hazardous objects, namely ‘knife’, ‘scissor’, ‘cup’ and ‘gas kettle’, to prevent child activities including those possibly dangerous objects. This was accomplished by training a YOLO model with image datasets and testing using another unseen set of images. As an additional feature the system is encapsulated with a mechanism that detects the danger level for the child by comparing the actual distance between the child and the object against a pre-defined distance of safety. A novel computer vision-based approach was implemented, combining, and attempting various innovative mechanisms to annotate bounding boxes focusing objects and children in order to calculate the proximity distance in between, to accomplish success in completing the specific objectives and delivering promising results. An alerting system was also integrated to trigger alarms towards the child and the parent according to the danger level detected. Models were trained, optimized, tested, and evaluated for numerous video frames representing different conditions of lighting, different angles and different object types, in order to create a system capable of preventing and minimizing the damage caused by hazardous objects.

8. REFERENCES

[1] "A home for paediatricians. A voice for children and youth."

[2] and S. A. R. Alya Al Rumhi, Huda Al Awisi, Mahmood Al Buwaiqi, "Home Accidents among Children: A Retrospective Study at a Tertiary Care Center in Oman," *Oman Med J*, [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6975254/#r5.

[3] S. Joseph, J. Ajay Gautham, A. Kumar, and M. K. Harish Babu, "IOT Based Baby Monitoring System Smart Cradle," *2021 7th Int. Conf. Adv. Comput. Commun. Syst. ICACCS 2021*, pp. 748–751, 2021, doi: 10.1109/ICACCS51430.2021.9442022.

[4] A. Rudyansyah, H. L. Hendric Spits Warnars, F. Lumban Gaol, and T. Matsuo, "A prototype of Baby Monitoring Use Raspberry Pi," *7th Int. Conf. ICT Smart Soc. AIoT Smart Soc. ICISS 2020 - Proceeding*, pp. 20–23, 2020, doi: 10.1109/ICISS50791.2020.9307586.

[5] M. Rai, A. Asim Husain, T. Maity, and R. Kumar Yadav, "Advance Intelligent Video Surveillance System (AIVSS): A Future Aspect," *Intell. Video Surveill.*, 2019, doi: 10.5772/intechopen.76444.

[6] V. Zeljkovic and D. Pokrajac, "Motion detection based multimedia supported intelligent video surveillance system," *Proc. Elmar - Int. Symp. Electron. Mar.*, no. June, pp. 49–52, 2006, doi: 10.1109/ELMAR.2006.329512.

[7] J. Lim, M. I. Al Jobayer, V. M. Baskaran, J. M. Lim, K. Wong, and J. See, "Gun detection in surveillance videos using deep neural networks," *2019 Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. APSIPA ASC 2019*, no. November, pp. 1998–2002, 2019, doi: 10.1109/APSIPAASC47483.2019.9023182.

[8] S. Loganathan, G. Kariyawasam, and P. Sumathipala, "Suspicious Activity Detection in Surveillance Footage," *2019 Int. Conf. Electr. Comput. Technol. Appl. ICECTA 2019*, pp. 2019–2022, 2019, doi: 10.1109/ICECTA48151.2019.8959600.

[9] "Video Surveillance For Elderly Monitoring And Safety."

[10] W. Y. Shieh and J. C. Huang, "Speedup the multi-camera video-surveillance system for elder falling detection," *Proc. - 2009 Int. Conf. Embed. Softw. Syst. ICESS 2009*, pp. 350–355, 2009, doi: 10.1109/ICESS.2009.62.

[11] V. C. Maha Vishnu, M. Rajalakshmi, and R. Nedunchezhian, "Intelligent traffic video surveillance and accident detection system with dynamic traffic signal control," *Cluster Comput.*, vol. 21, no. 1, pp. 135–147, 2018, doi: 10.1007/s10586-017-0974-5.

[12] J. Zhu, "Object Tracking in Structured Environments for Video Surveillance Applications," [Online]. Available: https://www.researchgate.net/publication/224586133_Object_Tracking_in_Structured_Environments_for_Video_Surveillance_Applications.

[13] G. Castañón, M. Elgharib, V. Saligrama, and P. M. Jodoin, "Retrieval in Long-Surveillance Videos Using User-Described Motion and Object Attributes," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 12, pp. 2313–2337, 2016, doi: 10.1109/TCSVT.2015.2473295.

[14] G. Percannella, D. Conte, P. Foggia, F. Tufano, and M. Vento, "A method for counting moving people in video surveillance videos," *EURASIP J. Adv. Signal Process.*, vol. 2010, 2010, doi: 10.1155/2010/231240.

[15] S. Vishwakarma and A. Agrawal, "A survey on activity recognition and behavior understanding in video surveillance," *Vis. Comput.*, vol. 29, no. 10, pp. 983–1009, 2013, doi: 10.1007/s00371-012-0752-6.

[16] D. Karthikeswaran, N. Sengottaiyan, and S. Anbukaruppusamy, "Video surveillance system against anti-terrorism by Using Adaptive Linear Activity Classification (ALAC) Technique," *J. Med. Syst.*, vol. 43, no. 8, 2019, doi: 10.1007/s10916-019-1394-2.

[17] K. Sitara and B. M. Mehtre, "Automated camera sabotage detection for enhancing video surveillance systems," *Multimed. Tools Appl.*, vol. 78, no. 5, pp. 5819–5841, 2019, doi: 10.1007/s11042-018-6165-4.

[18] B. Zhang, "Computer vision vs. human vision," no. 2004, pp. 3–3, 2010, doi: 10.1109/coginf.2010.5599750.

[19] S. Prasad and S. Sinha, "Real-time object detection and tracking in an unknown environment," *Proc. 2011 World Congr. Inf. Commun. Technol. WICT 2011*, pp. 1056–1061, 2011, doi: 10.1109/WICT.2011.6141394.

[20] H. Hashib, M. Leon, and A. M. Salaque, "Object Detection Based Security System Using Machine learning algorthim and Raspberry Pi," *5th Int. Conf. Comput. Commun. Chem. Mater. Electron. Eng. IC4ME2 2019*, pp. 11–12, 2019, doi: 10.1109/IC4ME247184.2019.9036531.

[21] E. Mühendisli, "Insansız Hava Araçlarının Derin Ö ˘ grenme Temelli Nesne Tespiti ve Tanıması Deep Learning Based Object Detection and Recognition of Unmanned Aerial Vehicles."

[22] A. R. Pathak, M. Pandey, and S. Rautaray, "Application of Deep Learning for Object Detection," *Procedia Comput. Sci.*, vol. 132, no. Iccids, pp. 1706–1717, 2018, doi: 10.1016/j.procs.2018.05.144.

[23] N. Yokoya and A. Iwasaki, "Generalized-hough-transform object detection using class-specific sparse representation for local-feature detection," *Int. Geosci. Remote Sens. Symp.*, vol. 2015-November, pp. 2852–2855, 2015, doi: 10.1109/IGARSS.2015.7326409.

[24] G. Chandan, A. Jain, H. Jain, and Mohana, "Real Time Object Detection and Tracking Using Deep Learning and OpenCV," *Proc. Int. Conf. Inven. Res. Comput. Appl. ICIRCA 2018*, no. Icirca, pp. 1305–1308, 2018, doi: 10.1109/ICIRCA.2018.8597266.

[25] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *Proc. IEEE Comput. Soc. Conf. Comput.*

Vis. Pattern Recognit., vol. 2016-Decem, pp. 779–788, 2016, doi: 10.1109/CVPR.2016.91.

[26] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, 2017, doi: 10.1109/TPAMI.2016.2577031.

[27] B. Liu, W. Zhao, and Q. Sun, “Study of object detection based on Faster R-CNN,” *Proc. - 2017 Chinese Autom. Congr. CAC 2017*, vol. 2017-January, pp. 6233–6236, 2017, doi: 10.1109/CAC.2017.8243900.

[28] W. Hongtao and Y. Xi, “Object Detection Method Based on Improved One-Stage Detector,” *Proc. - 2020 5th Int. Conf. Smart Grid Electr. Autom. ICSGEA 2020*, pp. 209–212, 2020, doi: 10.1109/ICSGEA51094.2020.00051.

[29] Y. Konishi, Y. Hanzawa, M. Kawade, and M. Hashimoto, “Fast 6D Pose Estimation Using Hierarchical Pose Trees,” *Eccv*, vol. 1, pp. 398–413, 2016, doi: 10.1007/978-3-319-46448-0.

[30] W. Liu *et al.*, “SSD: Single shot multibox detector,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9905 LNCS, pp. 21–37, 2016, doi: 10.1007/978-3-319-46448-0_2.

[31] H. Gao, “Understand Single Shot MultiBox Detector (SSD) and Implement It in Pytorch,” pp. 4–11, 2018.

[32] W. Zhu, N. Zeng, and N. Wang, “Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS® implementations.,” *Northeast SAS Users Gr. 2010 Heal. Care Life Sci.*, pp. 1–9, 2010.

9. APPENDICES

Final Paper - IT18006308

ORIGINALITY REPORT

18%	13%	8%	10%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Sri Lanka Institute of Information Technology Student Paper	3%
2	ai-pool.com Internet Source	1%
3	Submitted to University of Greenwich Student Paper	1%
4	dokumen.pub Internet Source	1%
5	towardsdatascience.com Internet Source	1%
6	viso.ai Internet Source	1%
7	lvivity.com Internet Source	1%

10	medium.com Internet Source	1 %
11	Guruh Fajar Shidik, Edi Noersasongko, Adhitya Nugraha, Pulung Nurtantio Andono, Jumanto Jumanto, Edi Jaya Kusuma. "A Systematic Review of Intelligence Video Surveillance: Trends, Techniques, Frameworks, and Datasets", IEEE Access, 2019 Publication	<1 %
12	Submitted to Sogang University Student Paper	<1 %
13	Chinmay Patil, Sumit Tanpure, Ankit Lohiya, Swaraj Pawar, Pratiksha Mohite. "Autonomous Amphibious Vehicle for Monitoring and Collecting Marine Debris", 2020 5th International Conference on Robotics and Automation Engineering (ICRAE), 2020 Publication	<1 %
14	Submitted to First City University College Student Paper	<1 %
15	Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi. "You Only Look Once: Unified, Real-Time Object Detection", 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016 Publication	<1 %

16	Submitted to Asia Pacific Instutute of Information Technology Student Paper	<1 %
17	Submitted to De Montfort University Student Paper	<1 %
18	Submitted to Bournemouth University Student Paper	<1 %
19	Submitted to University of London External System Student Paper	<1 %
20	Bin Liu, Wencang Zhao, Qiaoqiao Sun. "Study of object detection based on Faster R-CNN", 2017 Chinese Automation Congress (CAC), 2017 Publication	<1 %
21	hansheng0512.medium.com Internet Source	<1 %
22	Joseph K. Paul, Tankala Yuvaraj, Karthikay Gundepudi. "Demonstrating Low-Cost Unmanned Aerial Vehicle for anti-Poaching", 2020 IEEE 17th India Council International Conference (INDICON), 2020 Publication	<1 %
23	Submitted to Liverpool John Moores University Student Paper	<1 %

24	Submitted to University of Birmingham Student Paper	<1 %
25	"Smart Systems: Innovations in Computing", Springer Science and Business Media LLC, 2022 Publication	<1 %
26	Submitted to University of Hertfordshire Student Paper	<1 %
27	Meenu Ajith, Aswathy Rajendra Kurup. "Chapter 29 Pedestrian Detection: Performance Comparison Using Multiple Convolutional Neural Networks", Springer Science and Business Media LLC, 2018 Publication	<1 %
28	Submitted to Australian National University Student Paper	<1 %
29	Submitted to University of West London Student Paper	<1 %
30	N. Ananthakrishnan. "PROBLEMS AND LIMITATIONS WITH FINE NEEDLE ASPIRATION CYTOLOGY OF SOLITARY THYROID NODULES", ANZ Journal of Surgery, 1/1990 Publication	<1 %
31	ir.lib.uwo.ca Internet Source	<1 %
	Submitted to British University in Egypt	