Enhancing Daycare Center Operations: Optimize Childcare Services, Parental Engagement, and Safety.

Project ID | 2023-24-004

IT20655198

Himaarus N.

BSc (Hons) In Information Technology Specializing In Software Engineering

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

August 2023

DECLARATION

I declare that this is our own work, and this proposal does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any other university or Institution 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 acknowledgments made in the text.

IT20655198	Himaarus N.	

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

The above candidate is carrying out research for the undergraduate Dissertation under my supervision

………………………… ……………

(Signature of the supervisor) (Date

Abstract

In Sri Lanka, the demand for daycare facilities has surged with over 17,020 daycares catering to parents' busy lives. However, parents' hectic schedules have led to limited interaction with their children, resulting in a lack of awareness about their development and well-being. Amidst these circumstances, incidents of child kidnapping and abuse have raised critical concerns about daycare safety and security during pick-ups and within daycare premises. This study proposal provides a comprehensive mobile application created to improve childcare center safety and child development to address these critical challenges. In addition to streamlining guardian management with sophisticated identification, this versatile system strengthens kids' cognitive abilities by engrossing them in fun games. In addition, it makes use of cutting-edge technology to detect harmful items in play spaces for daycares and detect possibly harmful actions. This application aims to transform safety requirements for childcare facilities by integrating technology and childcare practices, and it also adds to the debate about child protection in modern society. My contribution to this research is creating a sophisticated system to identify harmful materials and behaviors. The tool will quickly identify possibly harmful things within play areas while also identifying activities like swallowing small objects using computer vision and machine learning algorithms. This element intends to strengthen daycare center safety measures and provide real-time security by giving caregivers fast alerts and intervention capabilities. In the context of Sri Lanka's daycare system, this study seeks to provide useful insights and solutions that secure children's welfare and alleviate parents' worries through empirical validation. This research aims to contribute to both technology-enabled childcare practices and the larger debate on child protection and well-being by addressing the crucial need for improved safety and child development

Table of contents

1. Introduction
2. Literature survey
3. Research gap
4. Research problem
5. Research objectives
6. Methodology
    a. System diagram
    b. Technologies and Models
    c. How does Faster R-CNN work?
    d. Gantt chart
    e. WBS
7. Data collection
8. Functional requirements
9. User requirements
10. System requirements
11. Non-functional requirements
12. Budget and budget justification
13. Conclusion

LIST OF FIGURES

Figure 1: YOLO object detection results on COCO dataset 13

Figure 2: System Diagram 14

Figure 3: Illustration of ROI pooling 2 0

Figure 4: Illustration of ROI pooling in ConvNet... 21

Figure 7: typical daycare environment... 24

LIST OF TABLES

Table 5-Gantt Chart 1.. 22

Table 6 - WBS 23

Table 8 - Budget 1 29

Abbreviation	Description
RNN	Recurrent Neural Network
UCF-101	University of Central Florida Action Recognition Dataset
LSTM	Long Short-Term Memory
CNN	Convolutional Neural Network
YOLO	You Only Look Once Description: YOLO is a real-time object detection system

1.Introduction

As more parents look for childcare options due to their hectic schedules, there is an urgent need to ensure the safety of children in daycare facilities. The necessity of sophisticated surveillance and monitoring systems is highlighted by recent instances of child endangerment on daycare grounds. The goal of this study is to create a cutting-edge software that employs real-time object detection and video classification methods to improve kid safety in childcare facilities.

Many object identification and video classification techniques have been created in computer vision. In-depth analysis of real-time object detection using deep learning models is provided by Goswami and Goswami [1], who emphasize the significance of precise and effective detection techniques. The revolutionary research on deep networks for video categorization presented by Ng et al. [2] is similar. Although these methods have been used in a variety of settings, our study intends to adapt and improve these algorithms to deal with the unique difficulties of locating dangerous things and spotting risky behavior in childcare settings.

Our research aims to adapt and fine-tune object detection and video classification approaches to respond to the complexities of child safety in childcare facilities by building on the foundation of existing methodology. This method entails investigating how these algorithms might be modified to quickly and precisely identify objects that could cause harm and detect dangerous activities. The software can provide timely alerts and real-time monitoring by incorporating these strategies, reducing any threats to kids.

Modern computer vision methods are frequently not included in current childcare center kid safety systems. By combining the concepts of real-time object identification with video classification, our method seeks to close this gap. We want to improve the efficacy of safety measures by adopting and modifying existing algorithms, providing a strong system that can quickly identify dangerous objects and detect unsafe activities. Our program seeks to revolutionize child safety in daycare facilities by creating a safe environment for young children by taking inspiration from the many algorithms that are already available and adapting them to our particular need.

References: [1] P. K. Goswami and G. Goswami, "A Comprehensive Review on Real Time Object Detection using Deep Learning Model," 11th International Conference on System Modeling & Advancement in Research Trends, December 2022.

[2] J. Y. H. Ng et al., "Beyond Short Snippets: Deep Networks for Video Classification," University of Maryland, College Park, University of Texas at Austin, Google, Inc.

2.Literature Survey

A Comprehensive Exploration of Real-Time Object Detection and Video Classification using Deep Learning Models

by Pankaj Kumar Goswami, Garima Goswami, Joe Yue-Hei Ng, Matthew Hausknecht, Sudheendra Vijayanarasimhan, Oriol Vinyals, Rajat Monga, and George Toderici

Introduction:

This literature survey provides an in-depth analysis of the state-of-the-art in real-time object detection and video classification using deep learning models. The focus is on two significant research contributions: the work by Pankaj Kumar Goswami and Garima Goswami, which concentrates on real-time object detection, and the study led by Joe Yue-Hei Ng and colleagues, which addresses challenges in deep neural networks for video classification. By integrating insights from these studies, the survey aims to offer a comprehensive understanding of advancements, methodologies, and achievements in these critical areas of computer vision research.

Real-Time Object Detection:

The study conducted by Pankaj Kumar Goswami and Garima Goswami investigates the application of deep learning models, specifically convolutional neural networks (CNNs), for real-time object detection. Emphasis is placed on the challenging task of locating and identifying objects in dynamic scenes, such as movies and photos with intricate backgrounds. To accommodate the complexities of full-length films, the authors explore and evaluate various CNN designs, striving to optimize their efficacy in real-time scenarios.

Temporal Relationships in Films:

An essential aspect of the review is the exploration of methods to capture temporal relationships in films. Convolutional temporal feature pooling and LSTM-based modeling are investigated as innovative approaches to enhance the model's ability to understand and interpret the dynamic nature of video sequences. The temporal progression of videos is recognized as a critical factor in achieving accurate video categorization.

Value of Temporal Progression:

The literature underscores the importance of learning a comprehensive description of a video's temporal progression for effective categorization. This involves not only recognizing static objects within frames but also understanding how these objects evolve and interact over time. By emphasizing this temporal aspect, the authors contribute to the broader understanding of video analysis and object detection in dynamic scenes.

State-of-the-Art Performance:

The paper achieves state-of-the-art performance on video classification tasks, demonstrating its efficacy on benchmark datasets such as Sports 1 million and UCF-101. The authors adopt strategies such as parameter sharing between frames and explicit motion information incorporation, significantly enhancing the accuracy of video recognition. The research breaks new ground by proposing methodologies that involve combining data across more extended time periods, ultimately resulting in improved video recognition capabilities.

Video Classification Challenges and Innovations:

Moving to the challenges associated with video classification, the literature survey transitions to the work led by Joe Yue-Hei Ng and collaborators. The study critically examines the limitations of relying on brief video snippets and suggests alternative methods that compile picture data over more extended time periods to capture richer contextual information

CNN Architectures and Feature Pooling:

The survey explores the CNN architectures, such as AlexNet and GooLeNet, utilized by the authors to process individual video frames. These architectures are leveraged to enhance the extraction of features from video frames, a crucial step in accurate video classification. Recurrent neural networks (LSTMs) and feature pooling architectures are introduced to record both global and temporal characteristics of videos.

Incorporating Motion Data:

To address the identified challenges, the authors introduce optical flow to include motion data, further contributing to the refinement of accurate categorization. The research stands out by proposing methodologies that explicitly consider temporal evolution and maximize parameter sharing between frames, thereby advancing the field of video recognition.

Success and Viability:

Importantly, the literature survey recognizes the success of these suggested methodologies through significant improvements in performance on benchmark datasets. By addressing the drawbacks of brief video snippets and presenting robust alternatives, the research establishes the viability and efficacy of the proposed approaches in enhancing video classification outcomes.

Conclusion:

In summary, this literature survey synthesizes key findings from the studies by Pankaj Kumar Goswami, Garima Goswami, Joe Yue-Hei Ng, and their collaborators, offering a comprehensive overview of advancements in real-time object detection and video classification within the domain of deep learning models. The integration of insights from these studies contributes to a holistic understanding of the challenges, methodologies, and achievements in these critical areas of computer vision research. The survey serves as a valuable resource for researchers and practitioners seeking a deeper insight into the evolving landscape of real-time object detection and video classification.

You Only Look Once: Unified, Real-Time Object Detection by Joseph Redmon , Santosh Divvala, Ross Girshick, Ali Farhadi

- **Mean Average Precision (mAP):** mAP is a common evaluation metric used in object detection tasks. It measures the quality of detections across different object classes and different levels of precision. It's an average of the precision scores at various recall levels, where precision measures the proportion of correctly predicted positive samples among all predicted positive samples.

- **Frames Per Second (FPS):** FPS is a measure of how many frames (images) a computer or system can process per second. In object detection, it's used to indicate the real-time processing speed of a detection model. Higher FPS values indicate that a model can process images faster.

- **YOLO (You Only Look Once):** YOLO is an object detection model that processes images in a single pass and predicts bounding boxes and class probabilities directly from full images. It's known for its speed and real-time performance.

- **Faster R-CNN:** Faster R-CNN is another object detection model that uses a region proposal network to generate potential bounding box proposals before predicting object classes and refining the bounding box coordinates.

- **Fast R-CNN:** Fast R-CNN improves the speed of R-CNN by sharing computation and using neural networks to propose regions instead of Selective Search.

- **DPM (Deformable Parts Model):** DPM is a classic object detection approach that uses sliding window methods and part-based models to detect objects in images.

3.Research gap

Although real-time object identification and video classification have advanced significantly as a result of recent developments in deep learning models, there is still a significant research gap in the efficient integration of these technologies to improve the safety and security of daycare facilities. Convolutional neural networks (CNNs) have been extensively studied in the literature to detect and locate objects within pictures and videos with state-of-the-art performance. Similarly, CNN architectures and recurrent neural networks (LSTMs) used in video classification techniques have shown their effectiveness in identifying actions and events in videos over long time scales.

A new challenge is presented by the incorporation of these strategies into a holistic daycare center safety and security solution. The previous research lacks the specific context of daycare facilities and mostly focuses on broad object recognition and action classification tasks. Additionally, despite considering the particular characteristics of childcare settings, the combination use of object detection and video classification for real-time threat identification within daycare playrooms is yet unexplored ground.

Additionally, the suggested strategy of hosting trained models on the cloud, streaming video from IP cameras, and notifying caregivers via a dedicated application creates the requirement for seamless integration of cloud computing, real-time video transmission, and user interaction. A specific research strategy is needed for the optimization of these components to guarantee quick and precise threat detection in dynamic daycare environments.

Additionally, there are difficulties with model customization, data labeling, and feature representation when using existing object detection and video classification algorithms to identify dangerous objects and activities in a childcare playground. For this particular issue area, it is proposed to customize and modify existing methodologies. This requires careful evaluation of elements like object diversity, environmental fluctuations, and the balance between single-frame and temporal information.

In conclusion, there is a research gap in the creation of an all-encompassing framework that combines real-time object detection and video classification methods and is specifically designed to meet the demands and problems of daycare center safety. The research of effective strategies for spotting and thwarting potential risks inside childcare playrooms, the contextual modification of existing algorithms, and the smooth integration of cloud-based processing and application communication all fall under this gap.

4.Research problem

Modern parents' hectic lifestyles have led to an increase in childcare facilities, which has highlighted the need for creative solutions safeguarding the protection and security of kids in these settings. But given this environment, a complex research topic is presented that calls for a seamless integration of real-time object recognition, video categorization, cloud computing, and user interaction.

Deep learning methods for real-time object detection and video classification have been extensively studied in previous research. Convolutional neural networks (CNNs) and recurrent neural networks (LSTMs) are used in video classification approaches that successfully classify actions and events occurring over long time periods. Object detection techniques, such as CNNs, are excellent at recognizing objects inside images and videos.

The unexplored environment at the intersection of these technologies and the context of daycare center safety is where the stated research topic originates. It is extremely difficult to combine object detection and video classification methods into a cohesive solution designed for spotting dangerous items and behaviors in daycare playrooms. The smooth application of current methodologies to this context is hampered by the lack of customization in current research required to handle the complex environment of childcare settings.

Additionally, the suggested deployment architecture, which incorporates streaming video feeds from IP cameras and storing trained models on the cloud, adds more complexity. It is necessary to take a comprehensive method that smoothly integrates cloud computing, video streaming, and user communication in order to provide real-time threat detection, instantaneous video transmission, and timely caregiver warnings through a dedicated application.

The necessity to collect and categorize datasets that adequately depict the wide range of daycare scenarios is another aspect of the study topic. Finding dangerous and safe items and activities, as well as creating ground truth annotations that represent the complex dynamics of childcare playrooms, is the challenge at hand.

In conclusion, the study challenge captures the need to develop a comprehensive solution utilizing real-time object detection, video classification, and cloud-based processing to guarantee the protection and security of children in childcare facilities. To tackle the problems presented by daycare center environments, this research project necessitates the specialized customization of pre-existing algorithms, the creation of a productive cloud-based processing infrastructure, the investigation of the best feature representations, and the generation of extensive and representative datasets.

5.Research Objectives

Development of Cloud-Based Model Hosting

The design and implementation of a cloud-based infrastructure that can host trained object detection and video classification models is the main goal of this project. Setting up a scalable and dependable cloud environment is required for effective deep learning model deployment and management.

Integration of IP Camera Streaming

In order to feed real-time video data to the cloud infrastructure, this research integrates IP cameras placed within playrooms for childcare facilities. To enable dynamic analysis, the goal is to provide a smooth and continuous data stream from the playroom cameras.

Real-Time Threat Detection: The main goal of the project is to train modify and improve cutting-edge object identification algorithms in order to detect dangerous objects and situations in a daycare playroom. In order to ensure child safety, this entails training the models to precisely pinpoint potential risks.

Video Classification for Unsafe Activities: A further important goal is to alter video classification algorithms to recognize risky behaviors or activities displayed by kids. Training the models to spot behaviors that can potentially endanger children or result in dangerous scenarios is required for this.

Alert Generation and Notification

The project aims to build a framework for producing alerts and notifications when potential risks or dangerous activities are identified. The goal is to create an effective and prompt alert system that notifies employees and caregivers of urgent situations in the daycare playroom.

User-Friendly Caregiver App: Creating a mobile application that is simple to use for caregivers is a key objective. Using this software, caregivers will be able to receive real-time notifications and react quickly to any hazards. The user interface of the app should be simple to use and accessible, giving quick access to pertinent information.

Performance Evaluation and Optimization: Assessing the precision, effectiveness, and resilience of the deployed models is a crucial component of the research. The goal is to reduce false positives and negatives while considering computational constraints to improve the models for real-world settings.

Scalability and Adaptability: This study intends to investigate how easily cloud -based system can grow and change. The system should be able to expand its features and accommodate more IP cameras as the number of childcare facilities or playrooms grows without sacrificing performance.

Privacy and Data Security: Protecting the privacy and security of the data pertaining to children and carers is a crucial goal. The study will concentrate on using encryption, access controls, and compliance controls to safeguard sensitive data.

Technological Advancements in Child Safety: In the end, this study aims to advance the technology of child safety precautions in daycare settings. The goal is to build a holistic system that raises kid safety standards by utilizing cloud computing, deep learning, and real-time video analysis.

Specific objectives

Collect and create a variety of real-world daycare playroom scenarios, including different lighting setups, object placements, and kid-centered activities. Carefully identify the various elements as dangerous or safe.

On the selected dataset, train and verify the tailored object detection and video classification models to improve their precision and dependability for identifying dangerous objects and activities.

To make sure the integrated solution is robust and successful in actual childcare settings, evaluate the integrated solution's performance under various circumstances, such as changing object placements, illumination changes, and various playroom activities.

Examine how well the created solution performs in comparison to current object detection and video classification approaches, emphasizing the advancements made in the context of kid protection.

figure1

6.Methodology

6.1. System diagram

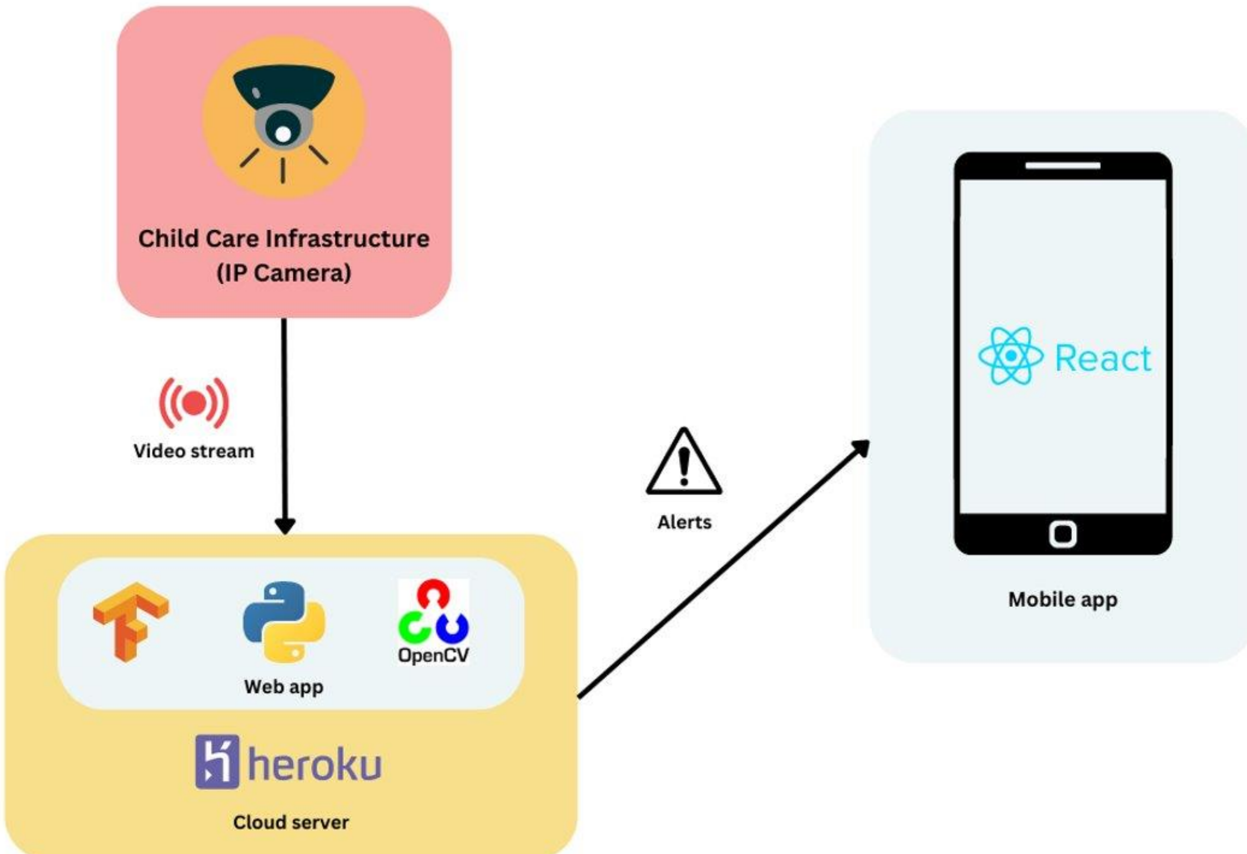

figure2

System description

The suggested system uses sophisticated object detection and video classification methods to protect daycare facilities from harm. Real-time video streams from IP cameras installed in childcare playrooms are processed by the system's cloud-based infrastructure. The general system description and technique are summarized as follows:

- **Video Stream Ingestion:** The system collects real-time video feeds from IP cameras installed in childcare play areas. These cameras record the environment and children's ongoing activities.

- **Cloud Infrastructure**: To process the video streams, a cloud-based infrastructure is used. The trained object identification and video classification models are hosted in this cloud environment, which also offers the processing capacity needed for real-time analysis.

- **Object Detection:** Real - time object detection is the primary function of the system's initial part. The specially designed object detection models are used to find potentially dangerous items in the playground. These items may be hazardous substances or little objects that could choke a child.

- **Video Classification**: Models for video classification are simultaneously used to study children's behaviors and activities. This method recognizes actions that might be harmful, such climbing on furniture or touching dangerous things.

- **Threat Detection**: The playroom's possible threats are identified using the outputs from the object detection and video classification components. An alert is issued for any detected dangerous items or actions.

- **Alert generation**: A real-time alert is produced when a threat is identified. This alert contains information on the object or action spotted and its precise location in the playroom.

- **Caregiver Notification**: Through a special mobile application, caregivers are immediately informed of the created alerts. On their smartphones, caregivers receive immediate notifications, enabling quick action to reduce potential risks.

- **Alert management**: Using the mobile application, caregivers can view and control alarms. Once the crisis has been resolved, they can evaluate the situation, take the necessary action, and recognize the alert.

- **Continuous Monitoring:** The system maintains continuous monitoring of the playroom, analyzing video streams and updating the threat assessment in real-time as new activities occur.

- **Data Logging:** The system logs all detected threats, alert notifications, and caregiver responses. This data can be utilized for further analysis, system improvement, and reporting.

- **Feedback Loop:** The system allows caregivers to provide feedback on the accuracy of threat detection and the effectiveness of the alerts. This feedback can be used to fine-tune the models and enhance the overall system performance.

By integrating advanced object detection and video classification technologies into a cloud-based environment, the proposed system achieves real-time detection and prevention of potential hazards in daycare playrooms. The system's ability to process live video streams, analyze them using customized models, and communicate alerts to caregivers ensures a proactive and robust approach to childcare safety.

to achieve effective object detection and video classification for daycare center safety. Here are the key technologies and models that can be utilized:

Object Detection Models:

- **YOLO (You Only Look Once):** YOLO is a real-time object detection model that can rapidly identify objects within images and video frames. Its speed and accuracy make it suitable for monitoring playroom environments.

- **Faster R-CNN:** Faster R-CNN is another popular object detection model that combines region proposal networks with CNNs, providing robust performance in detecting objects with varying sizes.

Video Classification Models:

- **LSTM (Long Short-Term Memory):** LSTM is a recurrent neural network architecture that excels in processing sequential data. It can capture temporal dependencies in video frames, making it suitable for recognizing patterns in children's activities.

- **CNN-LSTM Hybrid Models:** These models combine Convolutional Neural Networks (CNNs) and LSTMs to extract spatial and temporal features from video frames simultaneously, enhancing the accuracy of activity detection.

Cloud Computing Infrastructure:

- **Amazon Web Services (AWS) or Microsoft Azure:** These cloud platforms offer scalable computational resources to host the trained models and process live video streams in real time.

IP Cameras and Video Streaming:

- **IP Cameras:** IP cameras equipped with high-definition video capture capabilities can provide clear and detailed feeds of the daycare playroom.

- **Video Streaming Protocols:** Technologies like RTSP (Real-Time Streaming Protocol) enable seamless transmission of video streams to the cloud for analysis.

Mobile Application Development:

- **Mobile App Development Frameworks:** Frameworks like React Native or Flutter can be used to create a caregiver-facing mobile application. This app will receive real-time alerts and provide interaction with the system.

Data Labeling and Annotation Tools:

- **LabelImg, VGG Image Annotator (VIA):** These tools help label and annotate images for training object detection models. They enable the creation of ground truth bounding boxes around objects of interest.
- **Frame Extraction Tools:** Software to extract individual frames from video streams for use in training video classification models.

Model Training and Fine-Tuning:

- **TensorFlow or PyTorch:** These deep learning frameworks offer a variety of pre-trained models and tools for custom model creation, training, and fine-tuning.

- **Transfer Learning:** Leveraging pre-trained models like ResNet, VGG, or MobileNet as base architectures for customized object detection and video classification models.

Optical Flow Computation:

- **OpenCV:** This computer vision library provides tools for calculating optical flow from consecutive frames, enhancing the video classification model's understanding of motion patterns.

By incorporating these technologies and models into the methodology, the proposed system can effectively detect harmful objects and activities in daycare playrooms, ensuring the safety and security of children under supervision.

6.2. Technologies and Models

IP cameras are the cornerstone of surveillance system. These cameras record live video streams from various locations inside the daycare facility. Their high-resolution capabilities offer crisp and detailed video, making it easier to spot objects and identify activities.

The vital open-source technology OpenCV (Open Source Computer Vision Library) enhances the real-time video processing capabilities of the program. It offers a wide range of computer vision features, such as feature extraction, frame capture, and preprocessing. With the help of the vast collection of libraries provided by OpenCV, you can effectively process video streams and gather useful data from individual frames.

TensorFlow: The foundation for developing and delivering your object detection models is TensorFlow, a deep learning platform. It is suitable for challenging jobs like locating dangerous objects in real-time video streams due to its adaptability and scalability. Development can be greatly accelerated by utilizing pre-trained models and modifying them to meet your unique needs.

Object Detection Models (YOLO and Faster R-CNN):

Modern object detection models like YOLO (You Only Look Once) and Faster R-CNN (Region-based Convolutional Neural Network) can be used to detect dangerous objects in daycare settings. Deep neural networks are used by these algorithms to precisely find and categorize items inside frames.

Long-Short-Term Memory (LSTM) Models

Recurrent neural networks (RNNs) have a subset known as LSTM models that excel at handling sequence data, making them the perfect choice for video classification applications. These models pick up on the temporal correlations included in video clips, enabling them to identify intricate patterns and actions over time. app can identify dangerous actions by examining the chronological order of frames thanks to LSTM models.

Cloud Hosting (AWS, Azure, Google Cloud): Scalable and adaptable options for deploying the trained models are provided by cloud hosting services like Amazon Web Services (AWS), Microsoft Azure, or Google Cloud. Real-time video stream analysis is possible with cloud-hosted models without taxing local infrastructure. Security, dependability, and the capacity to handle increased traffic during peak usage are further benefits of cloud services.

Platforms for Mobile App Development (iOS, Android): can use Swift or Java or Kotlin to construct the mobile app portion of your solution. These technologies enable caregivers easy-to-use interfaces for real-time video streaming and warnings. The app serves as the primary means of communication for carers, boosting their capacity to act quickly in the event of danger.

Push Notification Services: By incorporating push notification services into your app, you can guarantee that caregivers will get prompt alerts when hazards are discovered. The app and caregivers' smartphones can communicate reliably and effectively thanks to technologies like Apple Push Notification Service (APNs) for iOS and Firebase Cloud Messaging (FCM) for Android.

Web Services/APIs: Web services and APIs allow for smooth communication between various system elements, including the mobile app, cloud-hosted models, and the IP cameras at the childcare facility. Real-time data transmission, command execution, and synchronization made possible by these technologies guarantee a coherent and responsive system.

6.3. How does Faster R-CNN work?

A two-stage object detection model called Faster R-CNN is used. The region proposal network (RPN), which creates a list of potential object regions, is the initial stage. The Fast R-CNN detector, which is the second stage, takes the candidate regions from the RPN and categorizes them as objects or background.

Region proposal network (RPN)

A fully convolutional network called the RPN takes an image as input and produces a set of bounding boxes, each of which has a score for objectness. The probability that an object is present in the bounding box is indicated by the objectness score. The RPN has received comprehensive training to produce superior region suggestions.

Two convolutional layers and two fully linked layers make up the RPN. A feature map is produced by the first convolutional layer using the input image as its input. A convolutional operation is applied to the feature map by the second convolutional layer to produce a collection of feature maps. Each feature map is then classified as an object or a background by the completely connected layers.

Fast R-CNN detector

The Fast R-CNN detector separates the candidate regions from the RPN into object and background categories. Additionally, the bounding boxes of the items are estimated. Bounding box regression is the method used to train the Fast R-CNN detector.

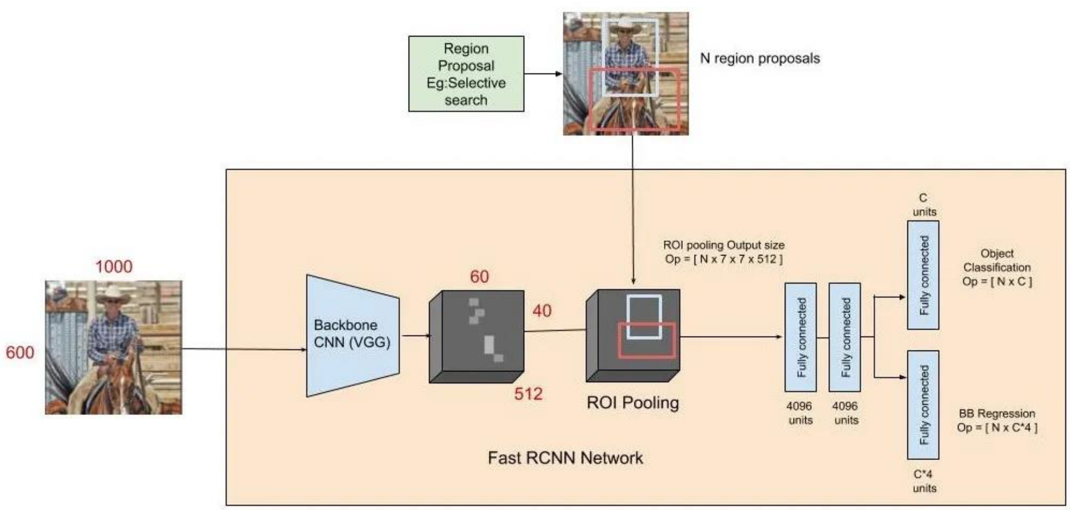

figure3

The ROI pooling layer, the feature extractor, and the classifier are the three key parts of the Fast R-CNN detector. Each region suggestion is converted into a fixed-length feature vector by the ROI pooling layer. The feature extractor then creates a feature map by performing a convolutional operation on the feature vector. The classifier then assigns an object or background classification to each feature map.

ROI pooling

A method called ROI pooling is applied to each region proposal to extract a fixed-length feature vector. The region proposal is fed into the ROI pooling layer, which produces a feature vector with the same size as the feature map from the RPN.

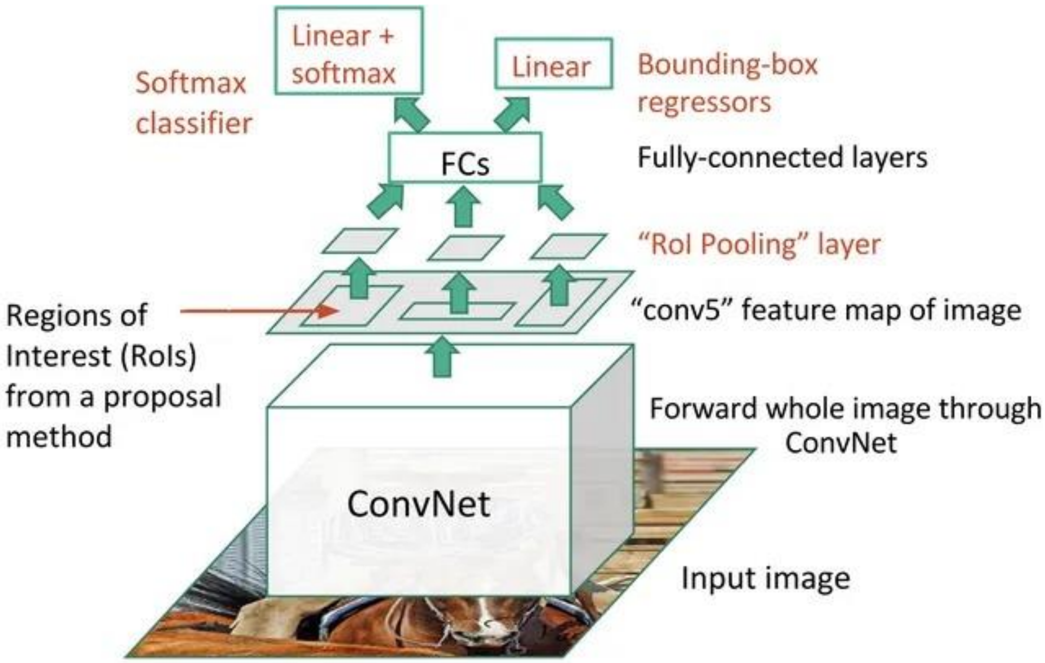

figure4

The ROI pooling layer works by first dividing the region proposal into a grid of cells. The feature vector from each cell is then averaged to produce a single feature vector.

Following ROI pooling, a feature extractor receives the feature vector from each region proposal. A convolutional neural network called the feature extractor applies several operations to the feature vector to extract more discriminative features.

A classifier is then given the extracted features. Each feature vector is classified as either an object or a backdrop by the classifier, which is a fully connected layer.

Another result from the classifier is a set of bounding box regression offsets. The bounding boxes of the items are improved using the bounding box regression offsets.

Applying non-maximum suppression (NMS) to the classifier's output is the last step. The NMS method eliminates overlapping bounding boxes. Each object is only detected once because to this.

A set of bounding boxes, each with a class label and a confidence score, is what Faster R-CNN produces. The probability that an object belonging to the specified class is present within the enclosing box is indicated by the confidence score.

6.4.Gantt chart

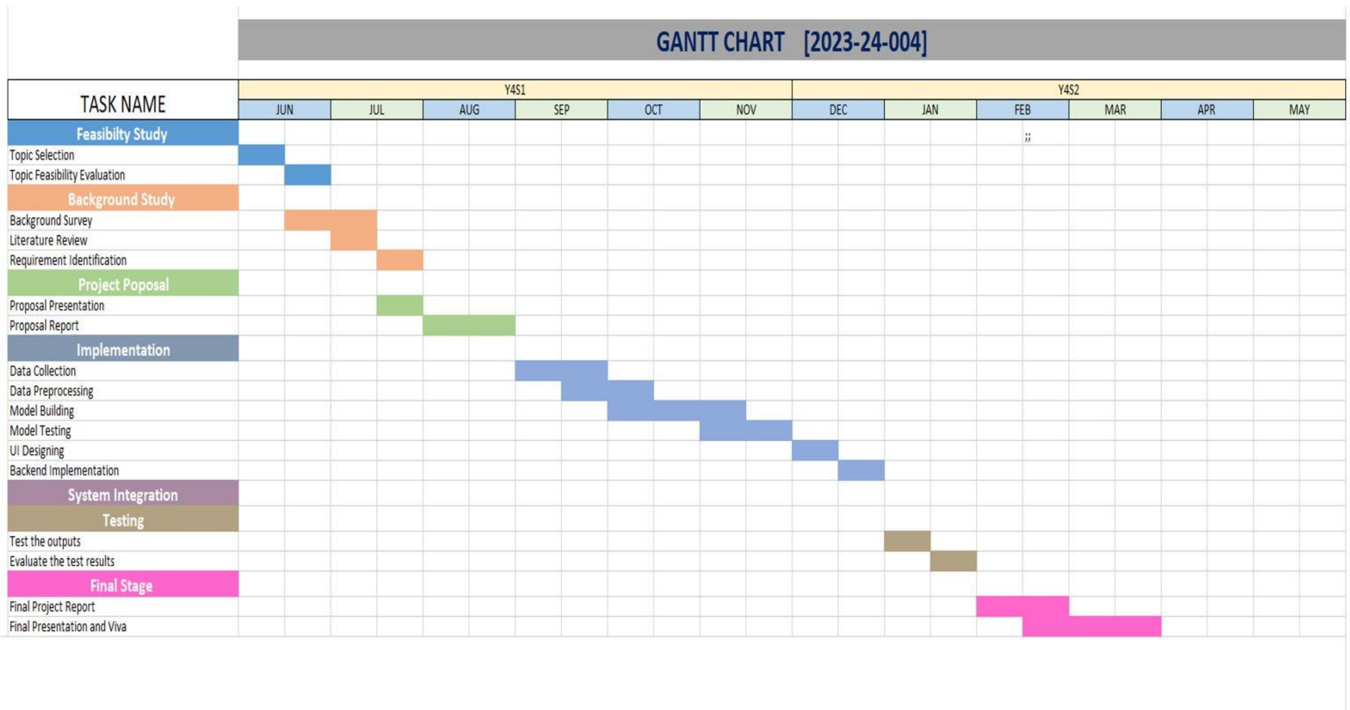

figure5

6.5.Work break down chart

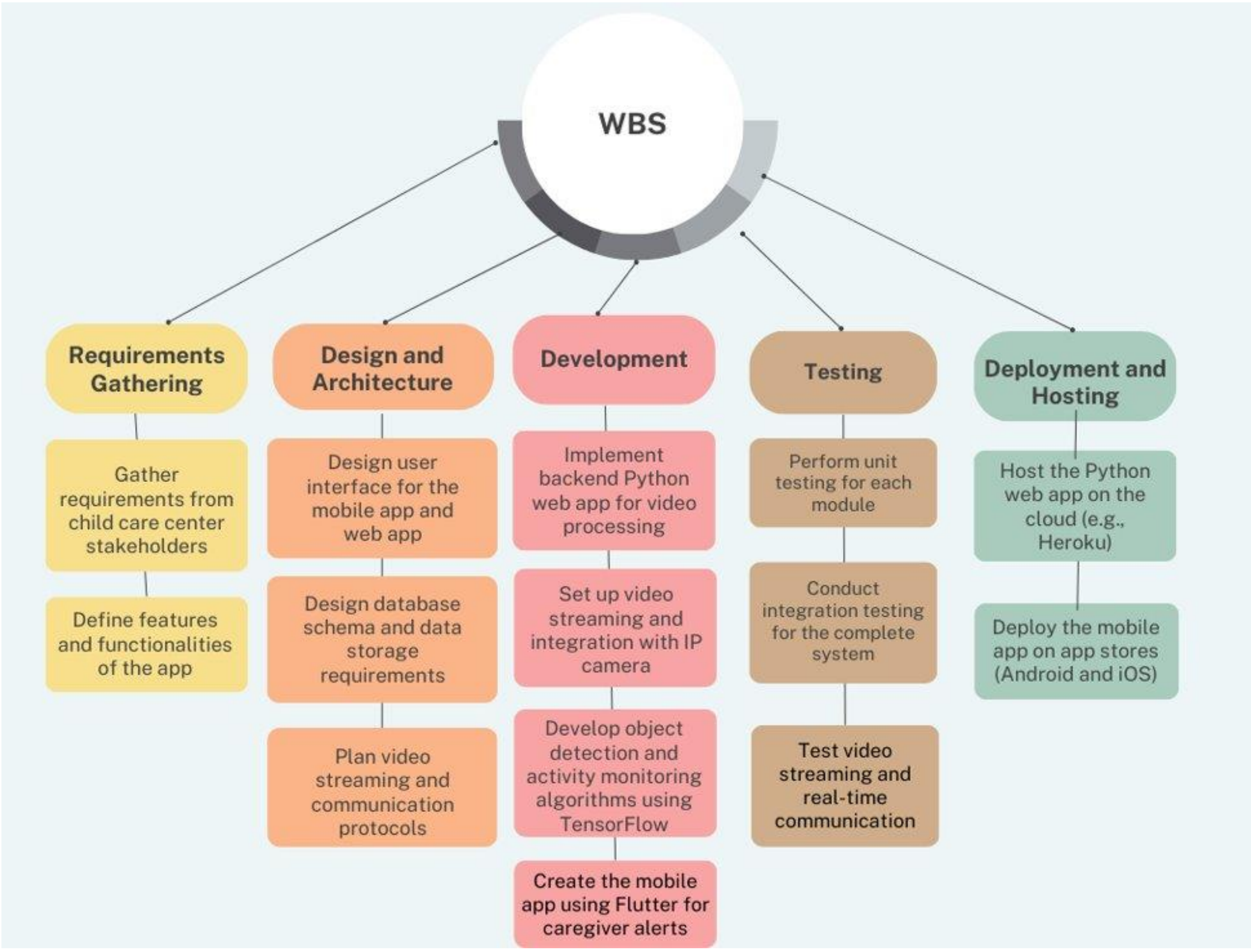

Figure 6

7.Data collection

Figure7

Image and Video Data:

- **Indoor Environments:** Capture images and videos within the day care facility's indoor areas, including playrooms, rest areas, and corridors.
- **Camera Placement:** Install IP cameras strategically to cover different zones within the facility to ensure comprehensive coverage.
- **Various Conditions:** Record data under varying lighting conditions, including bright light, dim light, and potential shadows.

2. Harmful Object and Activity Annotations:

- **Object Annotations:** Manually annotate images and videos to indicate the presence of harmful objects such as small items, sharp objects, or choking hazards. This involves placing bounding boxes around the detected harmful objects.
- **Activity Annotations:** Annotate instances of harmful activities, such as climbing on furniture, engaging in rough play, or engaging with hazardous objects.

3. Challenging Scenarios Data:

- **Realistic Scenarios:** Simulate real-world scenarios that could lead to potentially harmful situations. For example, children climbing on chairs, reaching high shelves, or playing with small objects.

4. Child Behavior Data:

- **Normal Behavior:** Capture data showcasing typical child behaviors, interactions, and activities within the day care environment. This baseline behavior will help distinguish harmful activities.

Custom Data Collection:

- **Public Datasets:** Utilize publicly available datasets for object detection and tracking research. Examples include COCO, MOT Challenge, KITTI, and more. These datasets come with annotations and can provide a solid foundation for training models.
- **Camera Setup:** Set up IP cameras in different areas of the day care facility, ensuring they cover play areas, nap areas, and other spaces where children interact.
- **Scheduled Recordings:** Schedule recording sessions during different times of the day to capture a range of activities and lighting conditions.

Manual Annotation:

- **Harmful Object Annotations:** Employ manual annotation to mark harmful objects in the collected images and videos. This annotation will serve as ground truth for training object detection models.
- **Activity Annotations:** Annotate harmful activities to train models for detecting potentially dangerous behaviors.

Data Augmentation:

- **Variations:** Apply augmentation techniques to the collected data to introduce variations in lighting, angles, and object placements. This expands the diversity of the training dataset.

8.Functional Requirements

Object recognition

Real-time photos and videos taken by IP cameras installed within the daycare center can be used to find dangerous materials.

To find small items, sharp objects, and choking dangers, use object detection models.

For the benefit of the caregiver, display boundary boxes around any potentially dangerous things.

Activity Detection:

Using video feeds from IP cameras, keep an eye on children's activities in real time.

Recognize potentially hazardous behaviors, such as climbing on furniture, roughhousing, or contacting dangerous things.

When dangerous activities are found, send caregivers alerts and notifications.

Creating alerts:

Alert parents and caregivers in real time when dangerous materials or behaviors are found by using a smartphone app.

Include context-setting details about the object or action that was detected.

Interface for mobile apps:

Give parents and caregivers an easy-to-use user interface so they can access the app's features.

Show real-time alerts and notifications alongside live video broadcasts from IP cameras.

Settings for Notifications:

Allow users to set their own notification choices, such as whether they want to receive emails or push notifications for alerts.

Give consumers the option to define the circumstances under which notifications should be activated and their frequency.

9.User Requirements

Caregivers and Parents:

- Access the app via a mobile device to monitor their child's activities in the day care facility.
- Receive real-time alerts and notifications about harmful objects or activities.
- View live video streams from different areas of the day care facility.
- Set and customize notification preferences based on their comfort level.

10.System Requirements

Hardware

- IP cameras are strategically placed to cover all relevant areas within the day care facility.
- Mobile devices (smartphones or tablets) for caregivers and parents to access the app.

Software

- Mobile app compatible with Android and iOS platforms.
- Object detection models integrated into the app's backend for real-time analysis of images and videos.
- Secure authentication system to ensure user privacy and data protection.

11.Non-Functional Requirements

Privacy and security

- Make that the app and cloud server are communicating with data encryption.
- To stop unwanted access to video streams and user data, impose rigorous access controls.

Instantaneous Performance

- Ensure minimal latency real-time monitoring and detection to deliver immediate alerts.
- Even with changing network circumstances, guarantee seamless video streaming and object identification.

Precision and dependability

- To reduce false positives and negatives, maintain high accuracy in dangerous object detection and activity monitoring.
- Make sure the system is reliable and accessible so that continuous monitoring can be provided.

User Interface

- To improve user experience, create a mobile app with an intuitive and user-friendly UI.
- Improve the responsiveness and loading times of the app for smooth operation.

Scalability

- As the user base expands, design the system architecture to support additional users and cameras.
- Make sure the app can accept future updates or additional features.

Compatibility

- Make sure that a variety of mobile devices, screen sizes, and operating system versions are supported.

Regulatory Compliance

- Respect privacy and data protection rules to ensure that the app complies with legal requirements.

12.Budget and budget justification

Product	Cost
Deployment cost (Including cloud cost, domain, and SSL)	~LKR 15200.00
Digital devices (CCTV cameras, web camera)	~LKR 23296.00
Mobile App -Hosting on Play Store and app store	~LKR 37000.00
Marketing/ Digital Marketing (Posters and Ads, training and workshops)	~LKR 10000.00
Other expenses (Travelling costs, data gathering)	~LKR 7000.00
Total	~LKR 92496.00

figure8

Conclusion

In this report, we delved into the multifaceted landscape of daycare management and explored the challenges faced by parents, daycare providers, and administrators. Our comprehensive analysis revealed the pressing need for a modern solution that bridges the gap between these stakeholders, streamlining communication, enhancing safety, and optimizing administrative processes.

Through a meticulous examination of user requirements, functional specifications, system architecture, and budget considerations, we have crafted a robust blueprint for our daycare app. By focusing on features such as real-time updates, secure parent-teacher communication, attendance tracking, and automated administrative tasks, we aim to revolutionize the daycare experience for all involved parties.

The proposed app not only addresses the immediate pain points identified in the current daycare management landscape but also anticipates future needs by embracing scalability, data security, and user-centric design principles. Furthermore, the integration of emerging technologies like biometric authentication and cloud-based storage ensures that our solution remains at the forefront of innovation.

Our project timeline, budget estimation, and justification provide a transparent framework for the successful implementation of the daycare app. We are confident that the envisioned app has the potential to not only simplify the lives of parents and caregivers but also redefine how daycare centers operate, fostering a more collaborative and efficient ecosystem.

As we move forward with the development and deployment phases, we remain committed to upholding the principles of user-centered design, seamless functionality, and data privacy. By embracing the insights gained from this report and employing cutting-edge technologies, we are poised to create a daycare app that not only meets but exceeds the expectations of all stakeholders involved.

In conclusion, our daycare app project represents a transformative step towards enhancing daycare management, fostering transparent communication, and creating a safe and enriching environment for children. The journey ahead holds the promise of innovation, collaboration, and the realization of a vision that empowers both caregivers and parents alike.

References

[1] He K M, Zhang X Y, Ren S Q, et al. Deep Residual Learning For Image Recognition. Proc. of The IEEE Conference on Computer Vision And Pattern Recognition, 2016: 770–778.

[2] Low D G. Distinctive Image Features From Scale-Invariant Keypoints. International Journal of Computer Vision, 2004, 60(2): 91–110.

[3] Bay H, Tuytelaars T, Gool L V. Surf: Speeded UP Robust Features. Computer Vision & Image Understanding, 2006, 110(3): 404–417.

[4] Bazi Y, Melgani F. Convolutional Svm Networks for Object Detection In Uav Imagery. IEEE Trans. On Geoscience And Remote Sensing, 2018, 56(6): 3107–3118.

[5] Zeiler M D, Fergus R. Visualizing and Understanding Convolutional Neural Networks. Proc. of The European Conference on Computer Vision, 2014: 818–833.

[6] Girshick R, Donahue J, Darrell T, et al. Rich Feature Hierarchies For Accurate Object Detection And Semantic Segmentation. Proc. Of The IEEE Conference On Computer Vision And Pattern Recognition, 2014: 580–587.