Smart Home Monitoring System for Early Childhood Using Computer Vision Technology

Zainab[1], Jalil Akbarzai[1], Muhammad Qasim[1], Ilsa Rehman[1], Bilal Ur Rehman[1,*], Muhammad Amir[1], Humayun Shahid[2], Kifayat Ullah[1]

[1]Department of Electrical Engineering, Faculty of Electrical and Computer Engineering, University of Engineering and Technology, Peshawar, Pakistan. [2]Department of Telecommunication Engineering, University of Engineering and Technology, Taxila, Pakistan.

* Correspondence: Bilal Ur Rehman bur@uetpeshawar.edu.pk

Citation | Zainab, Akbarzai. J, Qasim. M, Rehman. I, Rehman. B. U, Amir. A, Shahid. H, Ullah. K, "Smart Home Monitoring System for Early Childhood Using Computer Vision Technology", IJIST, Vol. 07 Issue. 02 pp 1306-1319, June 2025

Received| June 05, 2025 **Revised|** June 27, 2025 **Accepted|** June 29, 2025 **Published|** June 30, 2025.

One of the most significant problems families face today is the proper handling of newborns; most parents can barely always keep a close eye on their babies. Baby monitors put the minds of many parents at ease by increasing the safety of their children; however, many currently available models lack certain features that should comply with safety regulations. This paper proposes an intelligent monitoring system for infants that can be integrated into smart homes to improve real-time monitoring through a computer vision technique. Therefore, the primary goal of a smart home presence detection system is to enhance children's safety by accurately identifying their presence and identifying risks that may arise in real-life scenarios. It operates in real-time to ensure parents are always informed of their child's safety. This approach employs YOLOv5, which is well-known, fast, and accurate, thus suitable for this task due to its impressive real-time object detection performance. The proposed system indicates a quick and efficient framework for keeping children secure in smart homes, presenting the potential of advanced computer vision techniques in the real world.

Keywords: Computer vision, Machine Learning, real-time detection.

Introduction:

The welfare of a child is always a primary concern for parents and caregivers. While traditional safety strategies remain important, it is equally vital to recognize the emerging safety opportunities offered by modern technologies, particularly advancements in computer vision. Computer vision, a subfield of artificial intelligence, allows computers to analyze visual data using various algorithms and machine-learning techniques. This technology can detect and assess potential threats in real time, offering continuous monitoring of a child's activities and alerting caregivers to any signs of possible danger. In each epoch, caretakers have tried to protect their progeny from risks such as diseases, criminals, and accidents. Historically, it entailed keeping children in places where parental control prevailed over the child. However, due to current technology, parenting has shifted a notch higher as parents protect their children [1].

Constantly observing a baby can be challenging, especially for working parents who may struggle to monitor their child at all times. This can create genuine difficulties in ensuring the child's safety and well-being throughout the day. The most common approach to ensuring a child's safety is to hire a caregiver for full-time supervision or enroll the child in a daycare center. However, both options can be disruptive and often leave parents with lingering concerns about whether their child is truly safe. In this context, baby monitoring devices offer a practical solution to alleviate parental anxiety and stress by providing continuous oversight and reassurance [2]. The YOLO object detection method has been established as one of the most efficient in realizing real-time operations while demonstrating excellent accuracy in recognizing specific objects. In other words, YOLO manages big datasets and reduces laborious preprocessing pipeline requirements, making it a better solution than traditional techniques. Moreover, system accuracy can be significantly enhanced by training models on application-specific datasets, as these datasets allow the model to better recognize and respond to relevant scenarios.

[3] In this research, we employed the YOLOv5 model, a YOLO architecture series model. The results exhibited by YOLOv5 show high accuracy and real-time performance with speeds up to 140 fps. Additionally, the newly introduced YOLOv5 model possesses a considerably smaller weight file: it is 90 % smaller compared to the YOLOv4 model [3]. This study necessitates a specialized dataset tailored for training and labeling, aimed at accurately classifying instances where children are present in potentially dangerous areas. The information within this dataset covers numerous scenarios and the distribution of demographics, which is an achievement in this field of study. Thus, the proposed system detects children in risky zones and zones close to fire or swimming pools and works in real mode. The outlined method possesses high accuracy and speed; therefore, the proposed material could be considered a primary tool for avoiding accidents and protecting children.

Research Problem:

Notwithstanding significant progress in home safety technologies, guaranteeing children's immediate safety continues to be a formidable task, especially in settings like swimming pools and kitchens where dangers are unanticipated. Conventional approaches, such as continuous oversight and physical barriers frequently prove ineffective owing to human constraints and variable enforcement. There is an increasing need for sophisticated, automated systems to proactively identify children's presence in hazardous areas and notify caregivers in real time. Computer vision (CV) technology provides an effective solution by allowing ongoing, precise monitoring without direct human involvement.

Related Work:

Numerous advanced systems have been developed to manage and improve child safety. One such example is enhanced noise reduction technology, which is specifically designed to monitor and regulate the infant's environment more effectively. This system changes light, temperature, and humidity components, effectively removing sound noise, thus making a suitable environment for the baby and relieving parents' burden to monitor these aspects [1]. A new system using Raspberry Pi and Pi cameras was developed to monitor a baby's motion and the accompanying audio, specifically during crying. This system provides real-time video of the baby's location so caregivers can do so from a distance [2][3]. Another approach involves the development of an automatic monitoring system that is microcontroller-based and specifically designed for infant care. This system aims to provide real-time observation and prompt alerts to enhance child safety. It operates by detecting a baby's cry as a trigger signal, prompting the cradle to automatically rock until the baby stops crying, thereby soothing the infant without requiring direct caregiver intervention. In addition, cameras are mounted on the cradle, whose output is designed to provide parents with improved confidence in the baby's surroundings [3][4]. Additionally, a monitoring system incorporating GSM technology has been developed to track vital parameters such as body temperature, moisture levels, pulse rate, and movement. The information is sent through a GSM network to the parents, keeping them in touch with their baby's condition [5]. Likewise, another infant monitoring system employs embedded devices and involves door, light, or voice-detecting elements and other sensor elements. The sensor data is displayed in the form of LEDs while an alarm sounds in case the caregivers need it [6].

An improved method employs IoT and deep learning models for monitoring an infant's physiological vital signs, including their heart rate, breathing rate, body temperature, and position while sleeping. The system also takes into account various environmental factors such as the presence of sharp objects, light intensity, noise levels, humidity, and temperature. By monitoring both physiological and environmental conditions, it offers a comprehensive and integrated solution for ensuring infant safety and well-being [7]. Moreover, the 'Smart Cradle', an IoT-based baby monitoring system, has been developed to monitor further parameters like pulse and body temperature. The system uses sound sensors to detect when the baby is crying to promptly give feedback to the parents [8]. Numerous other systems are designed to safeguard the physical well-being of children. For instance, a smart home baby monitoring system has been developed that utilizes an action recognition model to detect potential risks based on the baby's movements and activities. This approach allows for early identification of hazardous situations, enhancing the overall safety of the child within the home environment [9]. Another IoT technology system informs parents about the security of their children and tracks where they are. It communicates with the parent via SMS alert and MMS, using a picture grabbed by the device's built-in camera, if any of the sensors shows unusual activity [10]. An innovative infant monitoring system implemented by Raspberry Pi includes live streaming of videos and audio, monitoring room temperature and humidity, and compatibility with the Arabic language [11].

Real-time health monitoring is also done for deaf parents using Android devices. This system uses sensors to detect the baby's body temperature, sound, and pulse rate, and it alarms its parents in cases where a malfunctioning device is identified [12]. A voice detection and Bluetooth-enabled system was developed to help prevent the tragic occurrence of children being accidentally locked inside vehicles. In such a scenario, the system detects the child's presence and promptly sends a notification to the parent's mobile device, enabling swift intervention to ensure the child's safety [13]. A new system for monitoring infants for deaf parents has been introduced to turn baby cries into specific alerts, eliminating the disadvantages of traditional tone-based systems [14]. The smart home management system has been designed to monitor the kids' activities, such as controlling TV watching and gaming [15]. This system uses specific sensors to record the length of these activities, helping the parent develop a balanced schedule for the child. Research has also improved the constraints of smart home gadgets to ensure that they would not compromise child safety [16]. In addition, an system has been created to protect children from possible incidents regarding stoves in their homes. For example, this system incorporates a camera that detects a child near a stove and sounds a bell so that the caretakers do not leave the child unattended, thus discouraging him or her from causing an accident [17].

The primary aim of this research is to create a smart home monitoring system that uses computer vision technology to detect the presence of children in proximity to dangerous places, such as swimming pools and fire zones, in real time and to provide prompt notifications to carers. Hence, it is accomplished by employing the YOLOv5 algorithm for effective object identification, utilizing a custom-labeled dataset for enhanced contextual relevance, and including alert mechanisms appropriate for smart home surroundings.

This research is unexplored due to the high-speed object detection model (YOLOv5s) with a context-specific application to enhance kid safety in indoor spaces and brilliant homes. In contrast to earlier systems that depend exclusively on environmental sensors or non-visual methods, our approach integrates real-time computer vision, annotated datasets specific to risk zones, and deep-learning inference to provide a scalable, efficient, and more precise end-to-end monitoring solution for detecting critical risk situations involving children.

Research Objectives:

The main objective of this work is to design and assess a smart home monitoring system that improves kid safety by applying a computer vision approach. In real-time, the system is designed to identify children's presence in perilous locations, like swimming pools and fire-risk regions. Primary goals encompass:

- To develop a real-time detection system for kid protection utilizing the YOLOv5 algorithm.
- To develop a custom-labeled dataset for identifying children in hazardous surroundings.
- To provide a system that delivers real-time notifications to carers, facilitating prompt intervention.
- To assess the efficacy of the suggested system in terms of detection accuracy, real-time efficiency, and system dependability.

Novelty Statement:

This study presents an innovative use of YOLOv5 for monitoring kid safety in smart homes. The suggested system employs computer vision for continuous and automatic monitoring of children's safety in dynamic situations, contrasting with standard safety techniques that depend on human supervision or stationary sensors. The innovation employs a lightweight, high-velocity object detection model (YOLOv5) specifically designed to identify children in real time while reducing computing expenses. Moreover, technology is engineered to produce instantaneous warnings, augment reaction times, and elevate overall kid safety in smart home environments.

Methodology:

This section delineates the suggested methodology for creating a smart home kid safety monitoring system, emphasizing the detection of kids in proximity to perilous areas such as fires and swimming pools.

Figure 1 illustrates that the system adheres to a systematic process consisting of data collection, data labeling, model training, testing, and output creation. A comprehensive and varied dataset is gathered from several scenarios to guarantee the model's resilience against varying risk conditions. During pre-processing, photos are annotated with the Roboflow tool to delineate pertinent elements, including children and hazardous areas. The annotated dataset is utilized to train the YOLOv5 algorithm, selected for its real-time object identification proficiency and rapid processing. After training, the model undergoes testing with novel picture samples to assess its accuracy and efficacy in identifying high-risk situations. Upon validation, the system functions in real-time, scrutinizing live video feeds to detect occurrences of a kid nearing a potentially dangerous zone. Upon identification, the system activates an alert visually or via a linked device facilitating prompt caregiver intervention. This comprehensive strategy seeks to automate danger identification in smart home settings and improve kid safety with less human oversight.

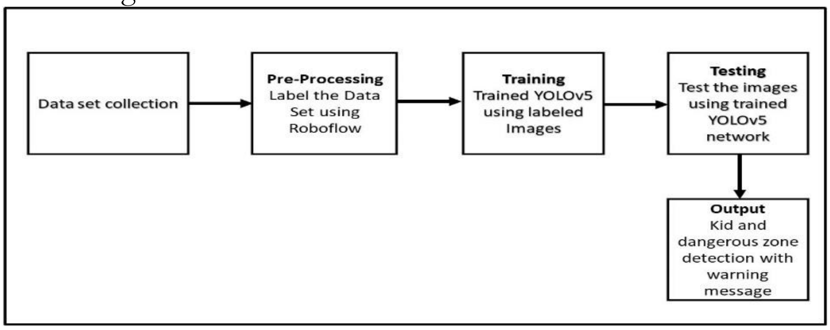

Figure 1. Flow diagram of the proposed methodology.

System Design:

The advanced home security system continuously captures real-time video footage through cameras and analyzes it using computer vision algorithms. These algorithms detect the presence of children and monitor their surroundings to identify any potential hazards, enabling timely alerts and preventive measures. The system activates an alert when detecting a kid in proximity to a dangerous zone, such as a fire or swimming pool. This alarm can be exhibited on a monitor, dispatched as a message via a mobile application, or communicated remotely to the parents, facilitating prompt response and augmenting kid safety.

Dataset:

The dataset used for this study comprises 7,800 images, enhanced through multiple rounds of data augmentation to improve the model's performance and generalization capabilities. Datasets for fire, pool, and child detection have been collected and integrated into a unified dataset. To enhance the dataset further, both images and videos were labeled using the Roboflow labeling tool, enabling more accurate training of the object detection model. Furthermore, the dataset was divided into three distinct categories: 80% for training, 15% for validation, and 5% for testing. This partitioning strategy, illustrated in Figure 2, ensures effective model learning, performance evaluation, and generalization assessment.

Proposed Algorithms:

Deep learning, a sped-up, detailed algorithm best suited for big data sets, is widely used for human identification tasks. Target identification techniques in this domain are often categorized into two classifications:

The first category consists of two-stage object detection techniques: R-CNN (Region-based Convolutional Neural Network), Fast R-CNN, and Faster R-CNN, which are discussed in this paper as part of the methodology. The second type includes single-stage object detection methods, such as YOLO. These methodologies afford accelerated detection rates while maintaining high accuracy and are, therefore, more appropriate for real-time detection tasks. YOLOV5

Object Detection Algorithm:

This research employs YOLOv5, a lightweight single-stage object detection model. Since the YOLO algorithms are pioneering as an option in object detection, they employed a regression-based approach that simplifies the network model and speeds up the detection process. This improvement makes YOLO highly efficient for applications that require real-time processing.

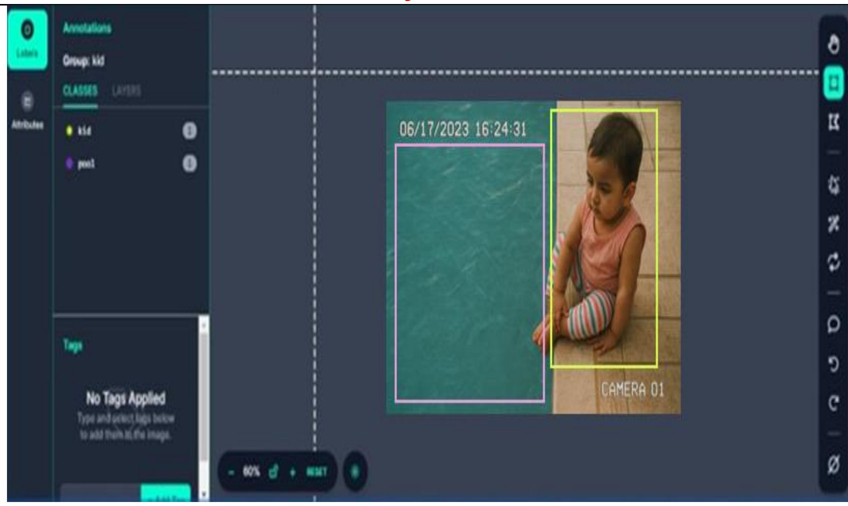

Figure 2. Roboflow labelling tool.

The architecture utilized in the proposed system, illustrated in Figure 3, comprises three primary components: the Backbone, Neck, and Head. The input images are processed in Backbone, which employs the convolution layers, using the C3_CBM blocks and BottleNeck to enhance the feature representation while obtaining a small image size. First, we leverage Concat and Upsample layers to stack those features and build a pyramid to detect objects of multiple scales better. Finally, the Head applies several Conv layers on the feature maps and generates detections at three different scales, namely 80×80×256, 40×40×512 and 20×20×1024, which supports M2SD. In the architecture, SPPF is used to aggregate contextual information into a single point for sure disposals, ensuring that detection is accurate regardless of the situation. This model shows higher performance in real-time child safety monitoring while maintaining a good balance of time and accuracy in defining a child's position in hazardous areas and providing instant alerts.

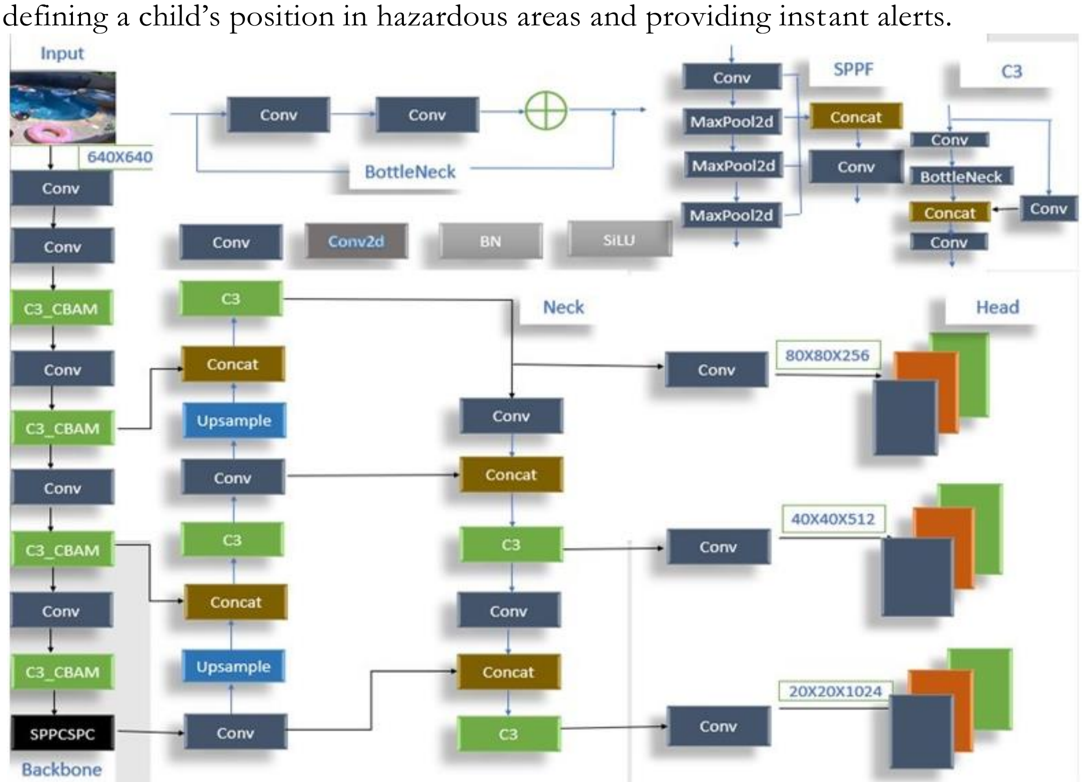

Figure 3. Architecture of YOLOv5 [18].

Figure 4. Detection of a child near a fire with an issued alert.

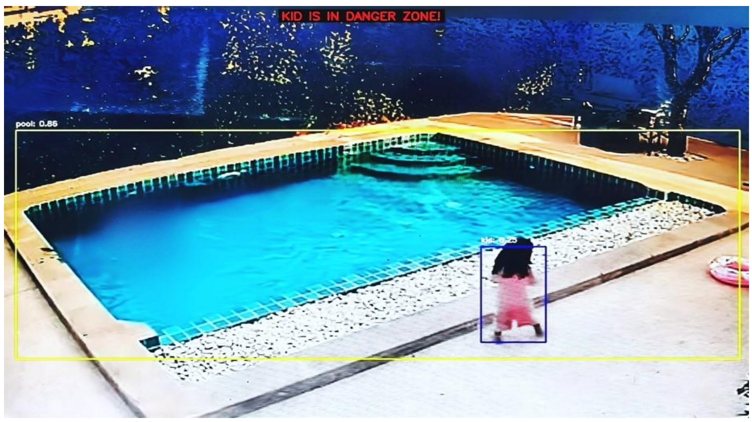

Figure 5. Detection of a child near a swimming pool with an issued alert.

Experimental Results:

This section presents the experimental results of the approach utilized for the child safety project. The system was evaluated in multiple scenarios involving situations where a child could be exposed to hazardous environments.

Scenario 1: Child near a fire with an issued alert.

Figure 4 presents a situation where a child is dangerously near the fire. The model effectively recognizes the fire and identifies the child, issuing an immediate warning to prevent potential injuries due to fire.

Scenario 2: Child near a swimming pool

Figure 5 demonstrates the scenario where the child is detected near a swimming pool. As demonstrated, the model accurately detects both the child and the pool as a hazardous zone. Once identified, it promptly triggers an alert, warning of the potential drowning risk.

Results and Discussion:

Tables 1 and 2 present the performance comparisons of different YOLOv5 models in detecting children, pools, and fire hazards, based on the outcomes of simulation experiments.

Each table evaluates the models using two specific metrics: Standard measurements such as Mean Average Precision at certain Intersection over Union (mAP for specific IoU values such as 50% (mAP50) and averaged over a more extensive range of IoU values (mAP50- 95). Table 1 also shows that the YOLOv5s version, which operates using the fastest detection mode, achieves the highest average performance. It records a MAP50 of 90%, significantly outperforming its YOLOv5l and YOLOv5m counterparts in terms of accuracy and efficiency. The proposed model gives very high accuracy in recognizing children (93.6% accuracy) and pools (96%), which are more important in enhancing the protection of children in unsafe environments. Further, a high fire detection accuracy of 80.4% indicates a reasonable level of capacity to identify possible fire threats.

Similarly, the YOLOv5m version demonstrates strong performance, achieving an overall MAP50 of 83.5%. It delivers particularly high detection accuracy for children at 85% and for pools at 93.8%, highlighting its effectiveness in identifying critical safety risks. However, the specificity of fire detection decreased slightly to 71.7%. The YOLOv5l model exhibits an equally good MAP50 of 81.6% for the chosen classes. Nevertheless, it is slightly slower in accuracy concerning the children, averaging 79.5%, and concerning fire hazards, an average of 73.3% compared to other models. This also raises concerns about the efficiency of the YOLOv5l model in basic real-time applications, particularly in scenarios where rapid detection of a child is crucial for ensuring safety.

Table 1. Simulation Results MAP50

YOLOV5 Version	Overall	KID	POOL	FIRE
YOLOv5l	81.6%	79.5%	92%	73.3%
YOLOv5m	83.5%	85%	93.8%	71.7%
YOLOv5s	90%	93.6%	96%	80.4%

The findings shown in Table 2 are dissimilar to the previous trend: YOLOv5sl, YOLOv5m, and YOLOv5s reported significantly lower MAP50-95 markers in comparison to the primary form; however, YOLOv5s remains the best at 68.4%. As shown in this model, compared with other variants, the recognition accuracy of children (73.1%) and pools (83%) has been improved; however, its fire detection precision (48.9%) raises high challenges for it to recognize fire threats stably.

The YOLOv5m model has a MAP50-95 of 60.8 %, with MAP50 scores near that: 60.8% for children and 78.7% for pools. However, it records a reduced fire detection precision of 42.9%, suggesting potential challenges in complex fire detection scenarios where greater accuracy is required. On the other hand, the proposed YOLOv5sl model gives the lowest MAP50-95 scores for all classes, particularly for detecting children (55.8%) and fire hazards (44.2%). This underperformance shows the limitation of this model in field application since time and accurate alerts are critical in protecting children.

Table 2. Simulation Results MAP50-95

YOLOV5 Version	Overall	KID	POOL	FIRE
YOLOv5sl	59.2%	55.8%	77.75%	44.2%
YOLOv5m	60.8%	60.8%	78.7%	42.9%
YOLOv5s	68.4%	73.1%	83%	48.9%

Figure 7 presents the F1-Confidence curve, which demonstrates the correlation between the confidence threshold and the F1 score across various object categories: fire, kid, pool, and the summation of all classes. Moreover, the F1 score acts as a deciding rate, the average precision and recollecting rates, which abridges too many false favorable and adverse rates. A higher score at the F1 metric indicates a better ground truth to detection precision ratio at a certain level of confidence.

The plot presented here also illustrates that the system's detection performance varies across different classes, indicating that certain object types are identified with higher accuracy than others. Finally, the F1 score of all classes, represented by the green line for the "pool" class; it can be seen that the "pool" class has the best F1 score in varying degrees of confidence, implying that it is the most straightforward class to classify without omission or commission error. On the other hand, the "fire" class (blue line) displays F1 scores that are significantly lower than others, which means that the model struggles much more to identify the case of fire, presumably because shapes, forms, sizes, or visual appearances of fire are diverse in the offered dataset. In addition, the overall performance for all classes, implemented as a blue-colored line, shows that the system achieved an F1 score of 0.87 when the confidence level was set to 0.414.

This value represents a perfect setting of the precision weight and recall weight of objects of all categories in the system, which means that the system performs well in real-world scenarios where they must balance the confidence levels to cover all objective detection requirements. Figure 6 shows the Precision-Recall curve of each category of objects of interest: fire, kid, and pool, and the overall system mAP@0.5 of 0.901 for all objects of interest. Specifically, the "pool" class demonstrates the best-seen precision/recall trade-off curve with mAP equal to 0.968, indicating the high efficacy of the proposed model in detecting pool-related dangers. The proposed "kid" class is performing well; it attained a mAP of 0.936, confirming how the formulated model can accurately detect children in vulnerable regions. Nevertheless, the system records a mAP of 0.797 for the "fire" class, indicating reasonable accuracy in detecting fire occurrences. However, this can be problematic, as fires in real-world conditions can vary significantly in form, size, and characteristics, potentially challenging the model's ability to generalize across diverse scenarios. The slight increase in the precision and recall means that the model does well in levels one and two in achieving the balance between true positives and the number of false positives. However, further improvement could be required to improve the fire detection performance. These discoveries affirm the system's benefits as an effective means of generating real-time identification of dangers to enable child security in smart homes.

The confusion matrix presented in Figure 8 demonstrates the model's classification performance across four distinct categories: fire, kids, swimming pool, and behind. The diagonal predominance is relatively high, especially concerning the "kid" class, where 439 patterns have been correctly classified. However, there are definite cases where classifiers got wrong, where kids were misclassified as 'pool', which was noted at 74. The "fire" class had 220 correct responses, along with 39 misclassifications identified with the other classes. Likewise, the background data has significant misclassification issues: 59 wrong labels as 'fire' and 33 as 'kid'.

The model is capable of accurately identifying specific categories, such as children and swimming pools. However, the results also suggest that there is still room for improvement, particularly in enhancing the model's ability to distinguish fire from background elements more effectively. Improving these misclassifications might notably enhance the system's performance, mainly if applied to essential applications such as fire detection.

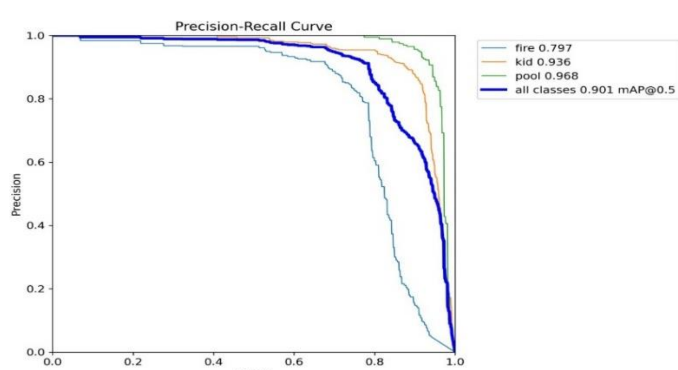

Figure 6. The P-R curve of the trained model.

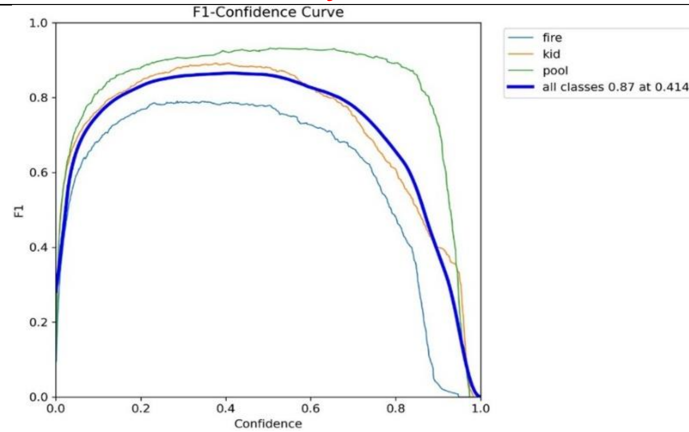

Figure 7. The P-R curve of the trained model.

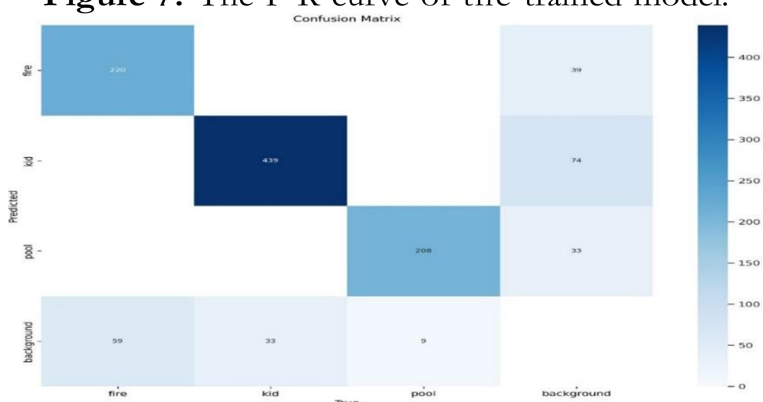

Figure 8. Confusion matrix of the trained model.

Discussion:

Our research illustrates the efficacy of the YOLOv5 model for real-time monitoring of kid safety in smart homes, attaining elevated detection accuracy and efficiency in perilous settings. Table 3 compares our model's performance with that of the current research.

Table 3. Comparison of Detection Accuracy and Real-Time Performance in Existing Studies

Literature	Detection Target	Accuracy (%)	Model Used	Real-Time Performance	Remarks
Current Study	Child detection near pools	93.60%	YOLOv5	140 fps	High speed and accuracy for real-time monitoring
[2]	Baby detection	85.00%	YOLO	Moderate accuracy	Slower performance in real-time applications
[3]	Crowd detection (similar to child detection)	87.00%	YOLOv3	30 fps	Lower accuracy and speed than YOLOv5
[4]	Baby detection in a cradle	80.00%	R-CNN	Slow (multiple stages)	High computational cost makes it unsuitable for real-time

[5]	Fire hazard detection	78.00%	CNN	Low fps	Difficulty in detecting fire hazards due to variability
[7]	Fire hazard detection	75.00%	Various object detection models	Moderate speed	Struggles with detecting fire hazards in dynamic environments

Detection Model Accuracy:

Our findings indicate that YOLOv5 attained a remarkable 93.6% accuracy in detecting youngsters around pools and 96% in recognizing the pools. The presented statistics exceed the identification rates documented by author[2] and author[3], who attained inferior accuracy levels of 85% and 87%, respectively, utilizing previous iterations of YOLO and other object detection methods. This indicates that the improvements in YOLOv5, especially its reduced model size and accelerated processing speed, boost its efficacy in real-time child protection applications.

Real-Time Performance:

The velocity of YOLOv5 (up to 140 fps) renders our system more appropriate for real-time monitoring compared to alternative systems, such as that created by author[4], which employed slower, multi-stage object identification algorithms like R-CNN. Although these models attain commendable accuracy, their computational expense renders them unfeasible for real-time applications. In contrast, YOLOv5 provides an optimal combination of rapidity and precision, rendering it suitable for scenarios where prompt reaction is essential, such as monitoring kid safety in dynamic settings.

Hazard Detection in Complex Environments:

Our system demonstrates robust performance in identifying children near pools, with 96% accuracy; however, the detection accuracy for fire risks is comparatively lower at 80.4%. Hence, it aligns with the findings of author[5], who also identified difficulties in recognizing fire threats owing to their unpredictable characteristics in real-world settings. Fires manifest in several forms, ranging from blazing flames to smoldering embers, potentially complicating detection algorithms. Future research may investigate the integration of hybrid models that merge YOLO with additional sensing technologies, such as thermal imaging or smoke detectors, to enhance fire detection precision.

Limitations and Future Work: A primary shortcoming of our existing approach is its comparatively diminished accuracy in identifying fire dangers relative to child or pool detection. This problem is not exclusive to our research since other publications have observed such difficulties [6]. Author[7] assert that improving the identification of complex and irregular objects, such as fire, necessitates a more specialized methodology, perhaps using multi-modal data (e.g., thermal or infrared sensing). Furthermore, we intend to augment our dataset to encompass a broader range of real-world circumstances, enhancing the system's generalizability.

Contribution to Smart Home Technologies: Our research enhances the existing literature on integrating computer vision inside innovative house ecosystems. In contrast to conventional approaches that typically depend on manual oversight or stationary sensors, our technology offers continuous, automated monitoring with exceptional precision. The capacity to identify threats in real-time and notify caregivers represents a substantial enhancement compared to earlier systems, which generally lack both real-time functionality and the use of sophisticated algorithms.

Conclusion:

The deployment of YOLOv5 represents a significant advancement since it demonstrates its ability to perform object recognition in real-time and specifically evaluates children's safety and real-time situations that may endanger their lives; the designed system ensures that caregivers are notified to intervene. Thus, using YOLOv5s is ideal because it yields a reasonable detection rate while bearing low computational demands. Using MAP50 and MAP50-95 metrics, YOLOv5s has proven to identify children in dangerous areas better than ever. In addition, the proposed approach improves security and helps to demonstrate how contemporary computer vision technologies work in practice. This solution provides a detailed framework to manage risks and prevent mishaps to keep children safe. Further refinements can be made regarding the system's mathematical accuracy and time of delivery for other future projects. It will also be possible to implement better versions of YOLO or even integrate hybrid models with other features of Artificial Intelligence technologies, which boost detection accuracy, specifically in complex environments.

References:

[1] S. Brangui, M. El Kihal, and Y. Salih-Alj, "An enhanced noise cancelling system for a comprehensive monitoring and control of baby environments," *Proc. 2015 Int. Conf. Electr. Inf. Technol. ICEIT 2015*, pp. 404–409, Jul. 2015, doi: 10.1109/EITECH.2015.7162982.

[2] A. F. Symon, N. Hassan, H. Rashid, I. U. Ahmed, and S. M. T. Reza, "Design and development of a smart baby monitoring system based on Raspberry Pi and Pi camera," *4th Int. Conf. Adv. Electr. Eng. ICAEE 2017*, vol. 2018-January, pp. 117–122, Jul. 2017, doi: 10.1109/ICAEE.2017.8255338.

[3] A. Ranjan, N. Pathare, S. Dhavale, and S. Kumar, "Performance Analysis of YOLO Algorithms for Real-Time Crowd Counting," *2022 2nd Asian Conf. Innov. Technol. ASIANCON 2022*, 2022, doi: 10.1109/ASIANCON55314.2022.9909018.

[4] R. M. K. Palaskar, Rachana Pandey, Shweta Telang, Ashwini Wagh, Akshada, "An Automatic Monitoring and Swing the Baby Cradle for Infant Care," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 4, no. 12, 2015, [Online]. Available: https://www.ijarcce.com/upload/2015/december-15/IJARCCE 42.pdf

[5] M. S. R. J. Mr.A.R. Telepatil, Prashant Bansilal Patel, Miss. S.S .Yajare, "Intelligent Baby Monitoring System," *Int. J. Res. Advent Technol.*, vol. 7, no. 6, pp. 191–194, 2019, doi: 10.32622/ijrat.76201981.

[6] G. K. K. Dive, "Design of embedded device for incubator for the monitoring of infants," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 3, no. 11, pp. 541–546, 2013.

[7] L. J. Lai, "An Intelligent Baby Care System Based On Iot and Deep Learning Techniques," *Int. J. Electron. Commun. Eng.*, vol. 12, no. 1, pp. 81–85, 2018, [Online]. Available: https://www.scribd.com/document/554900589/An-Intelligent-Baby-Care-System-Based-on-IoT-and-Deep-Learning-Techniques#:~:text=Deep Learning Techniques-,An intelligent baby care system is proposed that uses IoT,%2C noise%2C humidity and temperature.

[8] 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, Mar. 2021, doi: 10.1109/ICACCS51430.2021.9442022.

[9] V. Adewopo, N. Elsayed, and K. Anderson, "Baby Physical Safety Monitoring in Smart Home Using Action Recognition System," *Conf. Proc. - IEEE SOUTHEASTCON*, vol. 2023-April, pp. 142–149, 2023, doi: 10.1109/SOUTHEASTCON51012.2023.10115146.

[10] E. K. K. M Nandini Priyanka, S Murugan, K N H Srinivas, T D S Sarveswararao,

"Smart IOT Device for Child Safety and Tracking," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 8, 2019, [Online]. Available: https://www.ijitee.org/wp-content/uploads/papers/v8i8/H6836068819.pdf

[11] S. A. and S. M. Ibrahim, Dina M.Hammoudeh, Mohammed Ali A, "Raspberry Pi-Based Smart Infant Monitoring System," *Int. J. Eng. Res. Technol.*, vol. 12, no. 10, 2019, [Online]. Available: https://www.researchgate.net/profile/Dina-Hussein-4/publication/336989588_Raspberry_Pi-Based_Smart_Infant_Monitoring_System/links/5dbdc0234585151435e2177e/Raspberry-Pi-Based-Smart-Infant-Monitoring-System.pdf

[12] Y. K. Faruk Aktaş, Emre Kavus, "A Real-Time Infant Health Monitoring System for Hard of Hearing Parents by using Android-based Mobil Devices," *Istanbul Univ. - J. Electr. Electron. Eng.*, vol. 17, no. 1, pp. 3107–3112, 2017, [Online]. Available: https://www.researchgate.net/publication/315696973_A_Real-Time_Infant_Health_Monitoring_System_for_Hard_of_Hearing_Parents_by_using_Android-based_Mobil_Devices

[13] S. K. V Srivathsa Bharadwaj K S, Raashi A S, Prajwala A N, Shrihari Bhat, "Child Presence Detection and Alerting System in an Unmanned Car," *pices*, vol. 4, no. 11, pp. 284–288, 2021, [Online]. Available: http://pices-journal.com/ojs/index.php/pices/article/view/232

[14] A. Mohite and P. D. Jadhav, "Infant Monitoring System for Deaf Parents," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 7, no. 3, pp. 527–534, 2021, [Online]. Available: https://ijsrcseit.com/paper/CSEIT2173163.pdf

[15] R. Madani, B. Alturki, S. Reiff-Marganiec, and W. Alsafery, "My Smart Remote: A Smart Home Management Solution for Children," *1st Int. Conf. Comput. Appl. Inf. Secur. ICCAIS 2018*, Aug. 2018, doi: 10.1109/CAIS.2018.8442015.

[16] S. Sangal and R. Bathla, "Implementation of Restrictions in Smart Home Devices for Safety of Children," *2019 4th Int. Conf. Inf. Syst. Comput. Networks, ISCON 2019*, pp. 139–143, Nov. 2019, doi: 10.1109/ISCON47742.2019.9036218.

[17] S. Alghamdi, "Alarm System Designed to Protect Children from the Risk of Ovens in Smart Homes," *Int. J. Comput. Sci. Telecommun.*, vol. 10, no. 2, 2019, [Online]. Available: https://www.ijcst.org/Volume10/Issue2/p3_10_2.pdf

[18] L. L. Jian Ren, Zhijie Wang, Yifan Zhang, "YOLOv5-R: lightweight real-time detection based on improved YOLOv5," *J. Electron. Imaging*, vol. 31, no. 3, 2022, doi: https://doi.org/10.1117/1.JEI.31.3.033033.

Copyright © by authors and 50Sea. This work is licensed under Creative Commons Attribution 4.0 International License.