Smart Pool Monitoring System Using Yolo

Abhilash Saini[1], Abhishek[2], Kavana M R[3], Padmavathi[4], Rashmi P C[5]
[1] *Student, Dept. of Computer Science and Engineering, Vivekananda College of Engineering & Technology, Puttur, Karnataka, India*
[2] *Assistant Professor, Dept. of Computer Science and Engineering, Vivekananda College of Engineering & Technology, Puttur, Karnataka, India*

***Abstract*— The *Smart Pool Monitoring System Using YOLO* project is designed to improve safety in swimming pools by using smart technology. It uses a computer program called YOLO (You Only Look Once), a real-time object detection algorithm. We have fine-tuned it with a dataset. It is used to identify and locate human figures within each frame. The algorithm outputs bounding boxes and confidence scores, which help isolate humans from the background. This can quickly look at video and spot people in the water. If someone is in the pool at an unsafe time or shows signs of being in danger, the system can immediately send warning through a message or alert the concerned people. The system works by uploading the recorded video, breaking it into frames (images), and then checking each frame to see if a person is present or drowning. If it finds something unusual, like a person struggling in the water, it raises an alert. This project helps lifeguards and pool staff act faster in emergencies.**

I. INTRODUCTION

Smart Pool Monitoring System is a computer vision-based project to enhance safety in swimming pools by making use of real-time object detection. Drowning incidents during off-hours or when the lifeguards are not actively monitoring the pool present a great safety concern. In this regard, the present project puts forward a proactive solution that analyzes continuous video footage through the YOLO object detection model to detect and classify humans or any unusual objects in the pool area.

This system will work as an automated surveillance mechanism, estimating whether a person is in danger and sending out prompt alerts. Recognizing the presence and posture of individuals in and around the pool helps detect certain drowning scenarios and quickly respond to them. The motivation behind Smart Pool Monitoring is to integrate AI into everyday safety systems, reducing response time and saving lives, if possible. Such a solution bridges the gap between AI and real-world emergency prevention, making your swimming pool 'smarter' and 'safer' with technology.

II. LITERATURE SURVEY

[1] Lin et al. introduced the Microsoft COCO dataset, which revolutionized object detection with a large-scale and richly annotated image collection. Thousands of instances of humans and objects were captured in a wide range of conditions, including occlusion, complex backgrounds, and changes in lighting. COCO became a foundation for modern deep learning architectures, where models could be allowed to perform better generalization across real scenes. In the case of drowning detection and pool monitoring systems, the COCO dataset allows models like YOLO to detect humans with high accuracy, even in visually challenging aquatic environments. It forms the essential backbone in training systems that recognize people in different postures and partial submersions.

[2] Tian et al. proposed a novel concept for pedestrian detection, wherein they did not consider the human body as a single unit for detection; instead, they focused on strong part-based features. Their model thus relied on deep learning to identify distinctive body parts, thereby enhancing its accuracy when the subject was partially occluded. The ability to detect humans under partial visibility contributes.

[3] Redmon et al. proposed the YOLO framework, which marked an important milestone by reconceptualizing object detection into a single regression problem. Its exceptional real-time speed without compromising detection accuracy was realized due to the processing of the whole image in one pass. This fact made it very suitable for real-world applications that require rapid responses, such as swimming pool surveillance. Simultaneous detection

of multiple objects by this model allows for instantaneous recognition of a possible drowning incident. Its unified architecture seriously outperformed previous approaches that relied on multi-stage region proposals, forming the foundation for all later advances along the direction of efficient object detection.

[4] Zhang et al. conducted a thorough comparison among state-of-the-art algorithms for pedestrian detection, namely Fast R-CNN, SSD, and YOLO. Their side-by-side comparative performance analysis emphasized the trade-offs between architectural design choices with respect to detection accuracy and computational efficiency. The results of this study on real-time detection scenarios provided the relevant rationalization for selecting YOLO during pool monitoring. In these applications, where even a few-second delay might claim lives in safety-critical situations, the research of Zhang et al. further justifies the appropriate choice of YOLO due to its acceptable speed and reliability in detecting the presence of humans.

[5] Gulli et al. provided a practical and conceptual framework for the implementation of deep learning algorithms using the Keras library. The book covers the main areas that involve the use of CNN, data preprocessing, and augmentation techniques, which improve model performance. For the drowning detection projects, the knowledge of the CNN architecture and data optimization techniques from the authors is necessary to enhance the robustness of the model. In fact, training principles, fine-tuning, and validation in this work directly support the development of deep learning-based pool monitoring systems using YOLO.

[6] In this study, Redmon et al. advanced the YOLO framework by developing YOLOv3, which incorporated a deeper feature extractor (Darknet-53) and multi-scale feature detection. These innovations enabled the model to handle smaller objects and partially visible entities more effectively. YOLOv3's speed and accuracy make it ideal for real-time pool surveillance, wereswimmers' heads or limbs might only appear briefly. The paper also introduced improved bounding box predictions and class confidence metrics, establishing YOLOv3 as a benchmark for real-time detection tasks in video analytics.

[7] Goyal et al. presented a drowning detection model using the data of surveillance video in conjunction with machine learning algorithms, including SVMs and decision trees. For this purpose, their system analysed the motion trajectories to identify the abnormal or still movements of individuals in distress. Although the study was not based on YOLO, it highlighted behavioural analysis as an important building block in the development of automated drowning prevention. The integration of such a motion-based analysis with object recognition capabilities provided by YOLO could lead to better decision-making processes within automatic pool monitoring systems, enabling quicker and more accurate alerts.

[8] Nguyen et al. explored the challenges of detecting humans in poorly lit or visually degraded conditions. They combined image enhancement with deep learning-based detection to improve visibility and reliability in low-light settings. This research is particularly relevant for indoor pools or nighttime monitoring, where lighting can affect recognition performance. By applying similar enhancement techniques alongside YOLO, pool monitoring systems can maintain operational accuracy and responsiveness even under suboptimal lighting conditions.

[9] Khandait et al. developed a real-time drowning detection approach using video surveillance to identify irregular swimmer movements. Their model performed frame-by-frame analysis to distinguish between normal and distressed behaviors, relying on image segmentation and motion tracking. This research directly supports the goals of Smart Pool Monitoring Systems, demonstrating how early detection can trigger timely alerts and prevent fatalities. The methodology shows how integrating AI-based motion analysis with object detection frameworks can lead to practical, life-saving applications.

[10] Lu et al. presented an advanced human pose estimation model designed specifically for aquatic conditions. Their system was capable of identifying limb positions and postures despite water distortions. Incorporating pose estimation with YOLO detection enhances the ability to differentiate between normal swimming activity and potential drowning postures.

III. PROPOSED METHOD

The Smart Pool Monitoring System follows a structured pipeline to ensure efficient detection and timely alert generation. First, cameras continuously monitor the swimming pool and capture real-time video. The recorded video is then broken into individual frames, which are preprocessed through resizing and normalization. Each frame is analyzed using the YOLO object detection model to identify humans or other relevant objects, and if a person is detected in the pool during unsupervised hours or unusual behavior is observed, the system immediately flags the situation. Once a potential risk is identified, alerts are automatically sent to the concerned authorities via SMS or email, and an audible alarm can also be triggered for quick response. Additionally, the system logs the time and frame of every detection, enabling authorities to review, verify, and assess each event effectively.

IV. METHODOLOGY

The system captures live video from poolside cameras and converts it into frames for analysis. Each frame is preprocessed to enhance clarity and remove noise. The YOLO algorithm then detects and tracks people in the pool, identifying unusual movements or signs of distress that may indicate drowning. If a drowning situation is detected, the system instantly triggers alarms and email alerts while recording the event for later review and safety improvement.

V. SYSTEM ARCHITECTURE

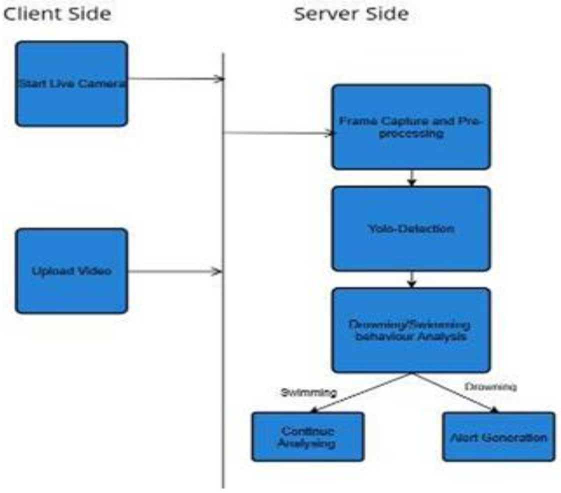

Figure 1: System Architecture

System architecture is the high-level conceptual framework that outlines how a system is organized, how it functions, and how its various parts interact. It provides a structured way to describe the system's components, their relationships, and their behaviour. An architectural description presents this information in a formal and organized manner, allowing developers and engineers to analyse, understand, and reason about the system's overall structure and operation.

1. Client Side
The client side serves as the entry point for video input. It provides two options for users to supply video data.

Start Live Camera: This option initiates a real-time video stream using a connected camera. It is suitable for continuous monitoring in environments such as swimming pools or open water areas.

Upload Video: This allows users to submit pre-recorded video files. It is useful for retrospective analysis or offline processing. Once the video input is selected, it is transmitted to the server for further processing.

2. Server Side
The server side performs the core computational tasks required for behaviour detection. It consists of the following sequential modules.

a. Frame Capture and Pre-processing
The incoming video stream is decomposed into individual frames. Each frame undergoes pre-processing operations such as resizing, normalization, and noise reduction. These steps clean and standardized for accurate analysis.

b. YOLO Detection
YOLO (You Only Look Once) is a real-time object detection algorithm. It is used to identify and localize human figures within each frame. The algorithm outputs bounding boxes and confidence scores, which help isolate swimmers from the background.

c. Drowning/Swimming Behaviour Analysis
This module analyses the detected individuals' movement patterns and postures. It uses spatial and

temporal features to classify behaviour into two categories:

Swimming: Characterized by rhythmic and coordinated movements indicating normal activity.

Drowning: Characterized by erratic, weak, or absent movements suggesting distress or danger.

d. Decision Flow
If the behaviour is classified as Swimming, the system continues to monitor and analyze subsequent frames. If the behaviour is classified as Drowning, the system proceeds to generate an alert.

3. Alert Generation
When drowning behaviour is detected, the system triggers an alert mechanism. Alerts may include visual notifications, audible alarms, or messages sent to emergency responders or monitoring personnel. This enables timely intervention and enhances safety in aquatic environments.

VI. FLOWCHART

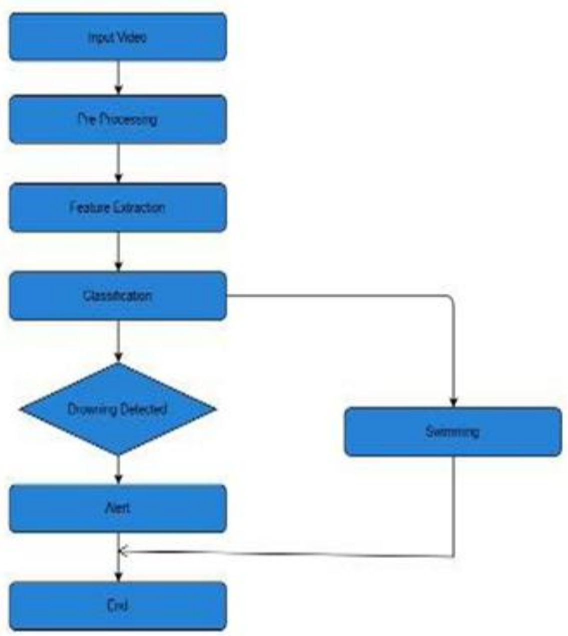

Figure 2: Flowchart Diagram

The flowchart of a video-based drowning detection system. The workflow starts with the input video, which first undergoes a pre-processing phase to improve visual quality and eliminate unnecessary noise. After this, the system carries out feature extraction, where key motion patterns and visual cues are captured from the video frames. These extracted features are then forwarded to the classification module, which evaluates whether the activity in the frame corresponds to normal swimming or a possible drowning scenario.

If normal activity is detected, the system simply continues monitoring. Conversely, if the classifier identifies signs of drowning, an alert is immediately generated to notify lifeguards or relevant authorities. After the alert is issued, the process concludes. This flowchart clearly presents the step-by-step functioning of an automated drowning detection system, ensuring quick identification and timely response to emergencies.

VI. CONCLUSION

The Smart Pool Monitoring System demonstrates how AI and computer vision can enhance pool safety. Using the YOLO model with OpenCV, Flask, and automated alerts, it analyses both live and recorded video to detect drowning in real time. Trained on a dataset of 5,000 images labelled as swimming or drowning, the system distinguishes normal movements from distress patterns.
Technologies like OpenCV for video handling, Flask for web monitoring, and SMTP for email alerts create a smooth, automated process from detection to notification. During testing, it showed high accuracy and quick alert generation through popups, sirens, and emails.
In summary, the project proves that AI-based monitoring can reduce human supervision, providing a reliable, continuous safety system that detects and responds swiftly to drowning incidents.

REFERENCES

[1] T. Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common Objects in Context," *Proc. European Conf. Computer Vision (ECCV)*, pp. 740–755, 2014.

[2] Y. Tian, P. Luo, X. Wang, and X. Tang, "Deep Learning Strong Parts for Pedestrian Detection," *Proc. IEEE Int. Conf. Computer Vision (ICCV)*, pp. 1904–1912, 2015.

[3] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, 2016.

[4] S. Zhang, R. Benenson, M. Omran, J. Hosang, and B. Schiele, “How Far are We from Solving Pedestrian Detection?,” *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 1259–1267, 2016.

[5] A. Gulli and S. Pal, *Deep Learning with Keras*, Birmingham, UK: Packt Publishing, 2017.

[6] J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” *arXiv preprint arXiv:1804.02767*, 2018.

[7] A. Goyal, M. Singh, and R. Sharma, “Machine Learning-Based Drowning Detection Using Surveillance Video,” *Int. J. Emerging Trends in Engineering Research*, vol. 8, no. 6, pp. 2301–2307, 2020.

[8] T. T. Nguyen, L. T. Nguyen, and T. V. Hoang, “Enhancing Human Detection in Low-Light Conditions Using Deep Learning and Image Processing,” *IEEE Access*, vol. 9, pp. 22345–22356, 2021.

[9] P. Khandait, R. Thakare, and S. Kale, “Real-Time Drowning Detection Using Video Surveillance,” *Proc. IEEE Int. Conf. Intelligent Systems and Applications*, pp. 450–454, 2021.

[10] H. Lu, X. Zhang, and Q. Wang, “Human Pose Estimation in Aquatic Environments for Drowning Detection,” *IEEE Trans. Image Processing*, vol. 30, pp. 4621–4633, 2021.