

## Literature review:

Vision-based systems and wearable sensor-based systems are two types of existing drowning detection technologies. Vision-based technologies are further subdivided into those that use underwater cameras [8], [9], [10], [11], [12], [13], [14] and those that use above-water cameras.

### Underwater

cameras have the drawback of missing the early struggle above the water. Early on, failure to recognize a drowning scene could result in a longer rescue time, which is a significant issue to consider in a time-critical emergency. The main disadvantage of a wearable-based system is the discomfort of use, which may lead to younger children seeking to alleviate the discomfort by removing the device, which is an unsubstantiated theory [15].

A. Object Detection Using Different Techniques .It is claimed that the usage of Convolutional Neural Network (CNN) architecture in Deep Neural Networks (DNNs) has added a significant shift in learning more complicated, informative characteristics in images as compared to the older techniques [16]. Furthermore, further optimized models such as Fast R-CNN, Faster R-CNN, and YOLO have been constructed since the region-based convolutional neural network (R-CNN) architecture proposal. Fast R-CNN, which improves bounding box (BB) regression and classification [17]; Faster R-CNN [18], which generates area suggestions using an extra sub-network [18]; and YOLO, which detects objects using a fixed-grid regression [19], are all faster than R-CNN. Bounding box regression is used to recognize generic objects based on basic CNN architectures. Local contrast enhancement and pixel-level segmentation, on the other hand, are used to recognize salient objects [20]. The techniques used in detecting objects under this chapter will be crucial as they establish the groundwork for the methodologies used to identify drowning and hazardous activities.

### B. Drowning Detection And Tracking

To avoid drowning events utilizing an alert system, Alshbatat et al. [10] proposed an integrated vision-based monitoring system consisting of a Raspberry Pi, two Pixy cameras, and an Arduino Nano board. They employed two cameras to detect and monitor swimmers by measuring their positions, and the swimmers were obliged to wear passive yellow vests. NEPTUNE [21], is another unique technology that uses statistical image processing [22] of video sequences to detect drowning victims as soon as possible. The equations utilized in detecting near-drowning victims are based on the variables created by statistical image processing. Another system called VIBE [15] uses background extraction to detect and track drowning victims and updates the motion area by taking the frame difference using the VIBE algorithm, which primarily evaluates the swimmers' positions when making judgments. How-Lung et al. [23] examine some difficulties in spotting drowning victims in a watery environment and offer an automatic detection surveillance system. The key obstacles in the aquatic environment, according to the authors, are water ripples and splashes, as well as background movements of the reflective zones. When it comes to recognizing swimmers, occlusions are also mentioned as a challenging difficulty. Their proposed solution is an algorithm that takes into

account all of these issues and detects water crises in complex aquatic environments [24], [25].

### C. Activity Detection Using Computer Vision

Current work on human motion prediction has been focused on two independent but complementary sub-tasks, according to Anand Gopalkrishnan [26]. 1) Short-term motion prediction, which is quantitatively evaluated by measuring the mean squared error (MSE) over a short period, and 2) long-term motion prediction, qualitatively evaluated by visual inspections of samples over a long period. Short-term models would be valuable in motion tracking applications because these jobs are applicable in several domains of work. On the other hand, long-term models might be valuable for creating computer graphic tools due to their broad applicability. Additionally, both models could be useful in human gait analysis, kinematics research, and human-computer interaction.

Existing drowning detection methods include vision-based systems and wearable sensor-based systems. Vision-based technologies are further classified as those that employ underwater cameras and those that employ above-water cameras. Underwater cameras have the disadvantage of missing the first battle above the water. Failure to detect a drowning incident early on may result in a lengthier rescue time, which is a key factor to consider in a time-critical emergency. The major downside of a wearable-based system is the pain of use, which may lead to younger children seeking relief by removing the device, which is an unfounded notion..

**Object Detection Techniques** - It is stated that the use of Convolutional NeuralNetwork (CNN) architecture in Deep Neural Networks (DNNs) has added a major shift in learning more sophisticated, useful features in pictures as compared to prior techniques. Furthermore, since the region-based convolutional neural network (R-CNN) architectural proposal, other optimised models such as Fast R-CNN, Faster R-CNN, and YOLO have been created. R-CNN is outperformed by Fast R-CNN, which enhances bounding box (BB) regression and classification; Faster R-CNN, which creates area recommendations using an additional sub-network; and YOLO, which identifies objects using a fixed-grid regression. Based on fundamental CNN architectures, bounding box regression is used to detect generic items.

**Object Detection Techniques** - It is claimed that the adoption of Convolutional NeuralNetwork (CNN) architecture in Deep Neural Networks (DNNs) has resulted in a significant improvement in learning more nuanced, usable characteristics in images as compared to previous techniques. Furthermore, since the region-based convolutional neural network (R-CNN) architecture concept, various optimised models have been developed, including Fast R-CNN, Faster R-CNN, and YOLO. Fast R-CNN, which improves bounding box (BB) regression and classification; Faster R-CNN, which generates area suggestions using an extra sub-network; and YOLO, which identifies objects using a fixed-grid regression, surpass R-CNN. Bounding box regression, which is based on core CNN architectures, is used to detect generic objects. How-Lung et al. - investigate several problems in detecting drowning victims in a watery environment and propose an automated detection surveillance system. According to the authors, the main barriers in the aquatic environment are water ripples and splashes, as well as background movement of the reflecting zones. Occlusions are cited as a major challenge when it comes to spotting swimmers. Their suggested solution is an algorithm that considers all of these factors and detects water emergencies in complicated aquatic habitats.

Computer Vision for Activity Detection According to Anand Gopalkrishnan, current work on human motion prediction has been focused on two distinct but complimentary sub-tasks. 1) Short-term motion prediction, which is tested quantitatively by calculating the mean squared error (MSE) over a short period, and 2) long-term motion prediction, which is evaluated subjectively by visual inspections of samples over a long period. Short-term models would be useful in motion tracking applications because these occupations are relevant in a variety of fields. Long-term models, on the other hand, may be useful for developing computer graphic tools due of their extensive applicability. Furthermore, both models may be beneficial in human gait analysis, kinematics research, and human-computer interaction..

Existing drowning detection technologies are split into two categories: vision-based detection systems and wearable sensor-based detection systems. Vision-based detection systems are further classified based on the location of picture acquisition: underwater cameras and overhead water cameras. Only vision-based detection systems are researched and analysed in depth in this section..

Zhang et al. suggested a camera-based drowning detection system that makes use of video sequences captured by underwater cameras. The background subtraction approach is used to detect a swimmer, which is then followed by an interframe-based denoising scheme to eliminate detection noise.

To reduce drowning accidents, Alshbatat et al. designed an automated vision-based surveillance system. A Raspberry Pi, two Pixy cameras, an Arduino Nano board, stepper motors, an alarm system, and motor drivers comprise their system. The swimmers were forced to wear passive yellow vests and were detected and tracked using two cameras that calculated their locations. The authors used innovative equations to introduce NEPTUNE, a near-drowning early prediction approach, in References. It utilises statistical image processing and k-means clustering to analyse picture frames to extract the drowning object's segmentation..

Wong et al. - used a thermal imaging system to identify moving human and water movement in a swimming pool throughout the day. This system is made up of two sub-algorithms that work together to identify an intruder both inside and outside the swimming pool by separating the pictures into two areas to accomplish the detection: head detection in both regions and water activity in just the second. Fei et al.- proposed a drowning detection system based on video frames captured by an underwater camera. To detect moving objects, the authors used the background subtraction approach. Fazanes et al. -recently examined and analysed the visual behaviour of drowning people utilising drowning movies.

Various approaches to intelligent swimming pool monitoring systems have been presented. The majority of approaches do background processing on input video frames. To detect a drowning person, some use background reduction and picture denoising. In, a Gaussian Mixture Model is used to describe the pixels, and the model's parameters are updated using the EM technique. Additionally, neural networks may be trained to distinguish between near-drowning and regular swimming movements. This, however, necessitates a huge dataset of both sets of behaviour. The dataset is collected by attaching a pressure

sensor to a swimmer who simulates drowning and regular swimming..

Swimmer detection methods can also benefit from pattern recognition techniques. In the backdrop, a model with past knowledge about swimming pools is used. This hierarchical model is based on behavioural patterns shared by nearly all disturbed swimmers. Another vision-based system is dependent on picture frame detection and intensity information. The YCbCr colour model is chosen for detecting water polo players in water, where brightness is separated and the Cb and Cr components are evaluated. Furthermore, underwater ultrasonic sensors detect drowning persons up to 70 metres below water in a swimming pool, and an underwater video detection unit locates and discovers the victims..

swimmers' body parts. This approach uses local motion

This study proposes a vision-based solution for recognising drowning people and alerting lifeguards in such cases. The individual swimming in the pool is recognised and tracked in this study utilising HSV colour space attributes and contour-based approaches. When the moving target is submerged for more than a predetermined amount of time, an alarm is sent to the lifeguard team. The HSV colour system was chosen over others because it is more successful at distinguishing the swimmer from the background in different lighting situations..