Spatial Information Research (2024) 32:143–150
https://doi.org/10.1007/s41324-023-00549-7

AI-driven drowned-detection system for rapid coastal rescue operations

Dileep P[1] · M. Durairaj[2] · Sharmila Subudhi[3] · V V R Maheswara Rao[4] · J. Jayanthi[5] · D Suganthi[6]

Received: 12 April 2023 / Revised: 23 August 2023 / Accepted: 28 August 2023 / Published online: 15 September 2023

Abstract

Recent observations indicate that nearly 50% of the public frequently visit coastal areas during weekends, seeking the health benefits of natural sunlight and fostering familial bonds. Notably, a significant portion of these visitors are unaware of swimming techniques or face other physical challenges, rendering them vulnerable to drowning, especially in areas lacking adequate lifeguard support or immediate medical emergency services. This study introduces an advanced drowned-detection device that employs a deep learning algorithm, grounded in artificial intelligence architecture, to swiftly detect and address potential drowning incidents. The system is particularly vigilant towards high-risk groups, such as children and the elderly. Upon detecting a threat, it autonomously deploys drones equipped with inflatable rescue tubes and notifies local authorities. Preliminary results suggest that our proposed model can effectively rescue a drowning individual in under 7 min, highlighting its prospective utility in curtailing swimming-related fatalities worldwide. This research underscores the need for technological intervention to enhance safety measures at coastal destinations and seeks to raise awareness about the importance of well-established lifeguard support.

Keywords Drowned-detection · Deep learning algorithm · Coastal safety · Autonomous drones · Swimming-related fatalities

1 Introduction

Drowning is a pressing global issue, particularly in children and young adults aged 1 to 14 years [1]. The World Healthcare Organization (WHO) ranks it as the third-leading cause of accidental deaths, especially alarming for children below the age of five. Annual drowning deaths worldwide are estimated at 236,000 [2]. With population growth and the rising popularity of swimming facilities in resorts and homes, this number is predicted to escalate. Current preventive measures, including public education, improved residential pool security, and increased surveillance at water bodies, fall

✉ Dileep P
dileep.p505@gmail.com

M. Durairaj
durairajtamili@gmail.com

Sharmila Subudhi
ssubudhi_cs@mscbuodisha.org

V V R Maheswara Rao
mahesh_vvr@yahoo.com

J. Jayanthi
jayanthij@sonatech.ac.in

D Suganthi
suganthiphd@gmail.com

1 Department of Computer Science and Engineering, Malla Reddy College of Engineering and Technology, Kompally, Hyderabad, Telangana 500100, India

2 Department of ECE, Peri Institute of Technology, Tambaram, Tamil Nadu, India

3 Department of Computer Science, Maharaja Sriram Chandra Bhanja Deo University, Baripada, Odisha 757001, India

4 Department of Computer Science and Engineering, Shri Vishnu Engineering College for Women (A), Bhimavaram, India

5 Department of Computer Science and Engineering, Sona college of Technology, Salem 636005, India

6 Department of Computer Science, Saveetha College of Liberal Arts and Sciences, SIMATS, Saveetha Nagar, Thandalam, Chennai 602105, India

short of adequately addressing the problem. This calls for a more effective solution: the use of automated, intelligent technology to combat drowning incidents [3–6].

Existing automated drowning detection methods can be broadly categorized into sensor-based and vision-based. Sensor-based methods require swimmers to wear equipment monitoring their physiological and environmental parameters [7]. Vision-based methods, in contrast, use cameras either above or underwater, coupled with deep learning techniques, to analyze swimming behavior. While these methods represent significant strides in drowning detection, they have limitations [8].

A comprehensive review of prior work reveals various approaches to automated drowning detection. Alotaibi proposed an IoT and transfer learning-based surveillance system using an overhead camera triggered by image sensors around the pool [9]. The system uses the ResNet50 network to categorize detected objects into three classes: person, animal, and item. Similarly, Li et al. developed a detection method using a modified version of the Yolov3 method, classifying individuals into four categories, with a reliability of 72.17% [10].

Alshbatat et al.'s proposal combines vision-based surveillance with hardware components, using two Squirty cameras, Raspberry Pi, and Arduino Nano boards to monitor and measure swimmer movements [11]. NEPTUNE and VIBE systems use analytical image analysis of surveillance videos and frame discrepancy calculations, respectively, to detect potential drowning. However, these methods also highlight the significant challenges in automated drowning detection, such as background movements of reflecting regions and water splashing, identified by How-Lung et al. [12].

The common theme across these studies is the importance of movement prediction, as discussed by Anand Gopalakrishnan. Recognizing movement patterns over both short and long terms is crucial to predicting and preventing potential drowning incidents, making it an integral aspect of any effective solution [13, 14].

Considering these findings, the current study aims to develop an innovative machine learning-based system that capitalizes on movement prediction capabilities to accurately identify potential drowning incidents. Using machine and deep learning algorithms, combined with automated devices like drones, this system goes beyond merely detecting incidents. It intervenes promptly by alerting authorities and initiating rescue operations, thereby enhancing the pool's safety and potentially saving lives.

The system utilizes deep learning technology, object recognition, and noise cancellation to monitor swimmers. It promptly identifies potential drowning incidents and alerts automated drones for a rescue operation, distinguishing itself from existing methods by its ability to recognize potentially dangerous activities and immediately notify the relevant authorities. The study aspires to prove that this technology can rescue a drowned person in less than 7 min. The work's objective is to pave the way for a new generation of intelligent, automated, and highly effective drowning detection systems that go beyond detection to proactive intervention, thereby addressing a significant global public health issue.

2 Materials and methods

Three key features of the technology described in this study are drone delivery to sufferers, detection of harmful actions, and drowning person detection. The drowning sensing element uses a customized LSTM prototype to identify sinking survivors. The prototype recognizes sinking in three phases and instantly alerts the consumer via sound. The second element seems to be the rescuing drones, which are launched in response to the signal for sinking identification and directed to the suspect's locations. The threat monitoring element will also use portable alerts to alert authorized staff on the property to possibly hazardous behaviors like racing around pools and consuming. It will encourage authorized workers to act responsibly, especially swimmers.

2.1 Drowning detection and tracking

Formation of the database: A database with 5000 photos was created because there was no previous data point for watery human organs. There are athletes in the sea in every single photograph in the database. Actual-time video recording and performer introductions are the main ways that information is gathered in terms of images. The Network was used as a secondary resource of information, and precise terms like "swimmer," swimming," drowning," as well as " sinking inside a pool" were used. Tagging of Images: The picture is marked using labeling, a Python-based visual application. Every picture is given a name by using configurable anchor boxes with YOLO-formatted predetermined tags. "Not drowning," Drowning step 1," Drowning step 2," as well as "Drowning step 3" seem to be the established keywords that are utilized in image captioning.

Initially, an aerial webcam is used to identify athletes in the water, as well as the DeepSORT method is then used to keep tabs on them. YOLO locates one or multiple things in the picture and sorts them to identify items. In comparison to other designs, Yolo functions well for a decent starting quality [20]. The majority of issues with detecting and tracking swimming athletes are dental, size alterations, then visual modifications. YOLO will be used to solve such issues [15].

A predetermined reference scheme has been used to determine the position of the athlete being monitored.

The water pool's specific specified reference frame. At first, swimmers are identified and their positions are monitored utilizing a predetermined reference scheme. Furthermore, it will look to see if the swimmers have reached any stages of drowning. As in case that such an occurrence is recognized, the webcam will monitor visual content in real-time and look for drowning casualties. The person recognition method will utilize the photographs with LSTM to determine who is drowning. Figure 1 shows the overview of the proposed methodology.

An auditory alert will ring to alert authorized employees of the incident after the geographical blocks of the drowned person have been appropriately pinpointed. For the graph to be recognized as a drowning scenario, the database's photos will now contain at least a single sinking individual. A drowning patient will be immediately identified by their position and motion. The "vertical ladders" workout, which mimics the motion of a human mounting a stair vertically [16], is among the most popular workouts despite being simple to detect.

2.2 Long short-term memory (LSTM)

The LSTM concept, which derives from the RNN framework of recurring neural networks, has been able to acquire long-term relationships, notably alterations in the concealed level. The forgetting gate, entry gate, as well as exit gate seems to be the three different kinds of gateways. The initial sigmoid level chooses which data to remove from cell states by analyzing the past result and present intake with in forgetting gate. The amount of data that has to be sent is determined using an integer that is between 0 and 1. No data is conveyed when the result is 0, but the entire data is delivered when the result equals 1. The input gateway decides whether or not the latest data will be stored. The sigmoid level selects which variable has to be updated similarly to the prior stage. The tanh level then generates fresh outcomes and updates the data by combining the data from both levels. The exit gateway then generates the exit signal using a filtering form.

$$i_t = \sigma\left(W_1^i \times X_t + W_h^i \times H_{t-1} + b_i\right) \, Inputting\, gate \tag{1}$$

$$f_t = \sigma\left(W_1^f \times X_t + W_h^f \times H_{t-1} + b_f\right) \, Forgetting\, gate \tag{2}$$

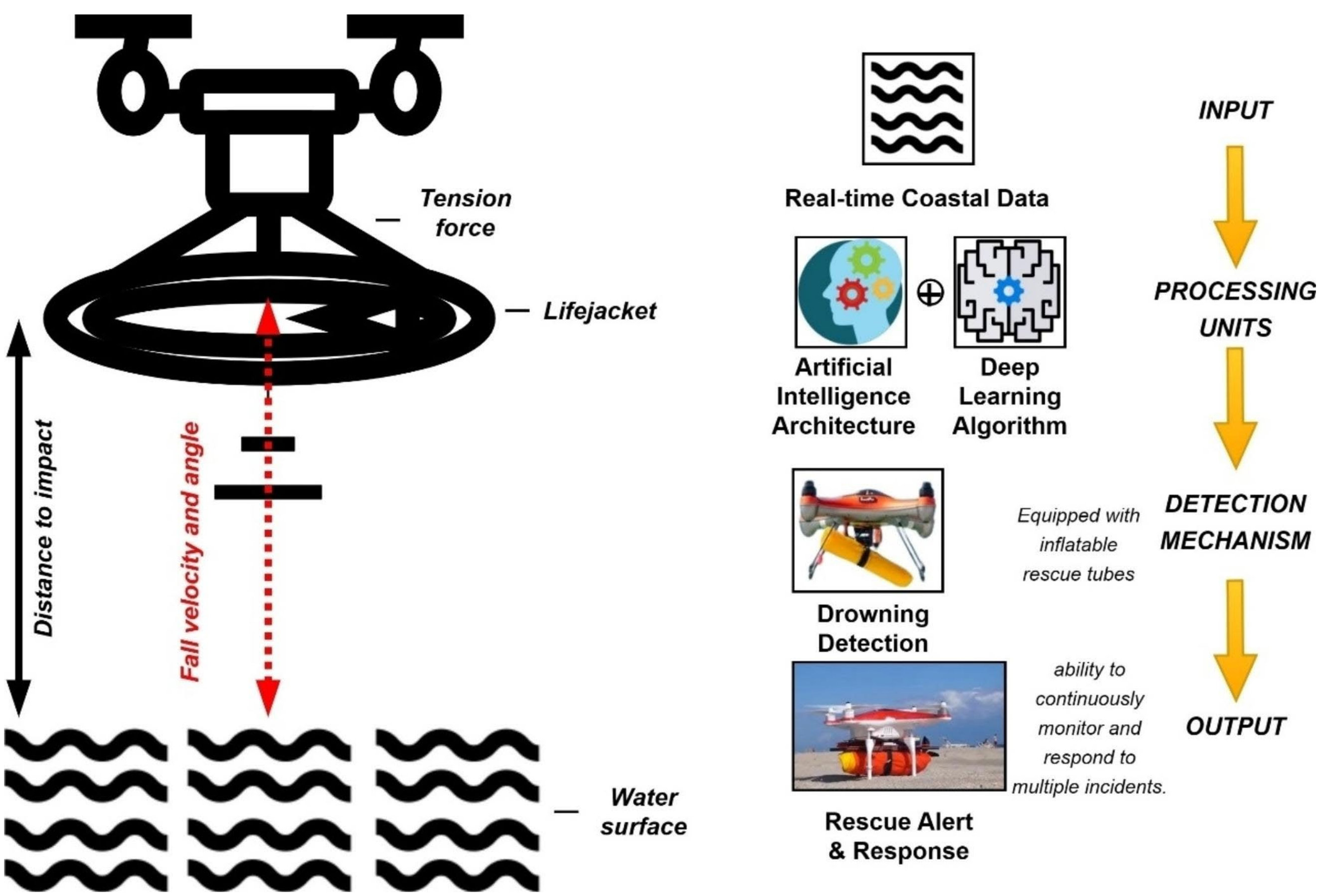

Fig. 1 Block diagram of the proposed methodology

$$o_t = \sigma\left(W_1^o \times X_t + W_h^o \times H_{t-1} + b_o\right) Outputting\, gate \quad (3)$$

$$\tilde{C}_t = tanh\left(W_1^C \times X_t + W_h^C \times H_{t-1} + b_C\right) Cell\, entrances \quad (4)$$

$$Sigmoid = \frac{1}{1 + e^{-1}} \quad (5)$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (6)$$

$$H_t = o_t \times \tanh\left(C_t\right) \quad (7)$$

While X_t is connected to such three gateways by weight matrices W_1^i, W_1^f, W_1^o and W_1^C, The weighted matrix connecting H_{t-1} to every gateway as well as cell entrances including W_h^i, W_h^f, W_h^o then W_h^C, b_i, b_f, b_o as well as b_C represent all gateways' as well as cells' entrances' biased factors, while $_\sigma$ represents the logistical sigmoid formula, m represents a vector-only characteristic in each gateway which receives data from features m of the such cell array, C_t represents the inner storage calculated in these units, Ht represents the result of a concealed phase obtained via storage multiplying, etc. The vanishing gradients issue with the RNN gets solved by the LSTM design. Figure 2 illustrates the algorithm for the proposed LSTM Classification method. C(t) = Cell state intrinsic storage, X(t) = Element-by-element entry, H(t) = The concealed state's outflow H(t1) = Earlier concealed state O(t) is the exit gate's component. i(t) = Input gateway component. Figure 3 shows the workflow of the LSTM in detection of drowning.

2.3 Identification of hazardous activities

To protect the security of the surfer, the device will also examine any apparent actions in the water. The presence of risky actions on the property is prevented through ongoing surveillance. It is done by sending portable alerts to authorized personnel informing them when a harmful action takes place. Gathering databases on harmful as well as non-toxic operations is the first step in the procedure of detecting harmful actions at a place. The database gets gathered in a pooling of individuals that are present in a social setting. Following the alert notice, the recognition procedure will be broken down into four stages:

2.3.1 Masking and noise extraction

The noises recorded from the location's seafloor are focused on wave motion throughout the daytime because of the cameras set up. To eliminate the chaotic parts of the picture, a basic stage of blurring the picture gets taken. Owing to each device's non-stationary behavior, each design's X, as well as Y parameters, serve as filtering spots.

2.3.2 Skeleton sketching

Actual-time multi-person vital feature recognition libraries [17, 18] are utilized to design skeleton-based people models to estimate a person's posture and then recognize that photograph in real-time [19]. Posture prediction effectively approximates a sophisticated nonlinear transformation matrix from a randomized picture of a human that matches the posture by combining DNN algorithms. Because of the fixed aspect of such webcam positioning, every picture has a resolution of 1280 × 720 pixels.

2.3.3 Labeling pictures using LSTM

A unique approach is developed using transferring training. The LSTM algorithm incorporates the drawn darkish back grounding visuals determined by the skeleton sketching procedure, determining the classes (toxic as well as non-hazardous actions). For every group designated for learning, the instruction is delivered in two phases.

2.3.4 Notifying the authorized personnel using mobile alerts

When frames get designated as dangerous, a state warning gets issued to any of the approved authorized gadgets. It will alert the relevant staff to start taking the required safety procedures and steps. Every Android mobile device will be able to utilize the Android application [20].

Step 1: Preprocessing

Step 2: Represent an element by ignoring bias.

Step 3: Creating W, representing the matrix of weight to transfer input from vectors of gate.

Step 4: Fine tuning.

Step 5: Build vector by applying tan function.

Fig. 2 Algorithm for the proposed LSTM Classification method

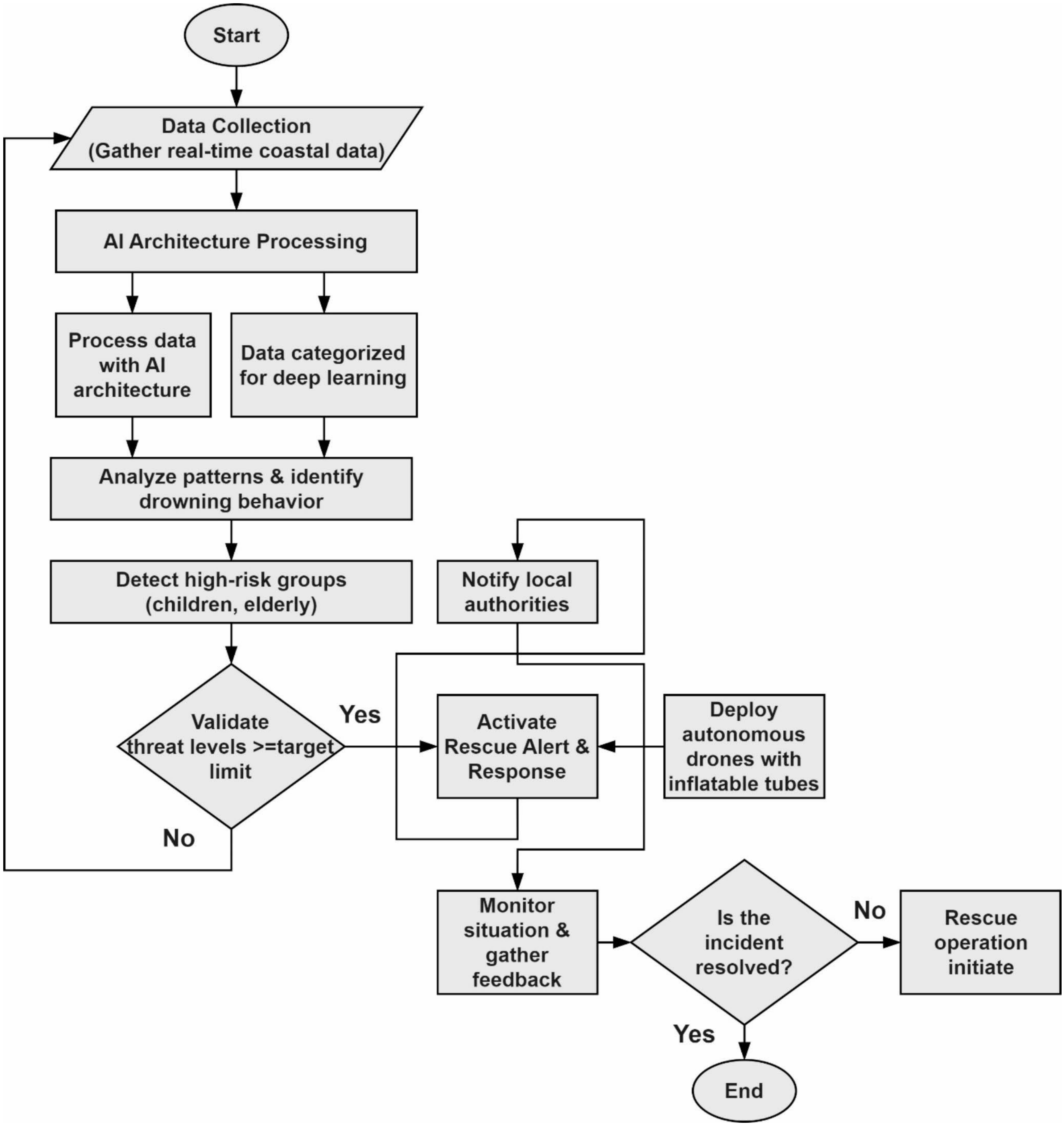

Fig. 3 Workflow of AI-Driven Drowned-Detection System

3 Result and discussion

Several performance indicators will be used to evaluate and contrast a machine learning algorithm with alternative approaches. Accuracy, sensitivities, specificity, as well as precision, constitute the most often used evaluative measures, along with F1 but also MCC. The proportion of accurately anticipated sinking as well as swimming incidents in the trial database is measured by the accuracy (Ac) statistic. Sensitivities or recall (R) seems to be the proportion of accurately anticipated drowning incidents compared to the overall sinking occurrences in the database. Precision (Pr) calculates the ratio of incidents of drowning that were correctly anticipated to all instances of drowning that have been forecasted. Specificity (Sp) seems to be the ratio of correctly anticipated non-drowning instances to every non-drowning event included in the database. The following algebraic representations of such measures are possible

$$Ac = \frac{(TP + TN)}{(TP + TN + FN + FP)} \tag{8}$$

$$R = \frac{TP}{(TP + FN)} \tag{9}$$

$$Pr = \frac{TP}{(TP + FP)} \tag{10}$$

$$Sp = \frac{TN}{(TN + FP)} \tag{11}$$

while, in turn, TP, TN, FP, as well as FN stand for examples of genuine positivity, true negativity, false positivity, and false negativity. An evaluating measurement called the F-measure (F1) combines retention and accuracy into a unique number. The Mathew correlate coefficient (MCC) seems to be a metric that achieves a balance between predicting sensitivities and selectivity. MCC seems to be a numerical measure with values ranging from -1, which indicates a perfect forecast, to $+1$, which indicates an inverted forecasting [21–23].

$$F1 = \frac{2PR}{P + R} \tag{12}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FN)(TP + FP)(TN + FP)(TN + FN)}} \tag{13}$$

The entry parameters for such YOLO detecting method [19] are 416×416. Employing a YOLO-dependent identification method, drowning fatalities get found in three phases. The DeepSORT method has been able to keep tabs on the swimmers while they were submerged for a considerable amount of time. The system's efficiency is shown in Fig. 4 plotting the false positive and true positive rates obtained by the proposed system for 500 images for different stages of classification of drowning.

Nonetheless, LSTM was trained across 15 epochs using two repetitions per epoch, which allowed the system to retrain and then verify the information incredibly effectively. Figure 5 shows the accuracy of the proposed methodology for different number of iterations. Over 30 repetitions,

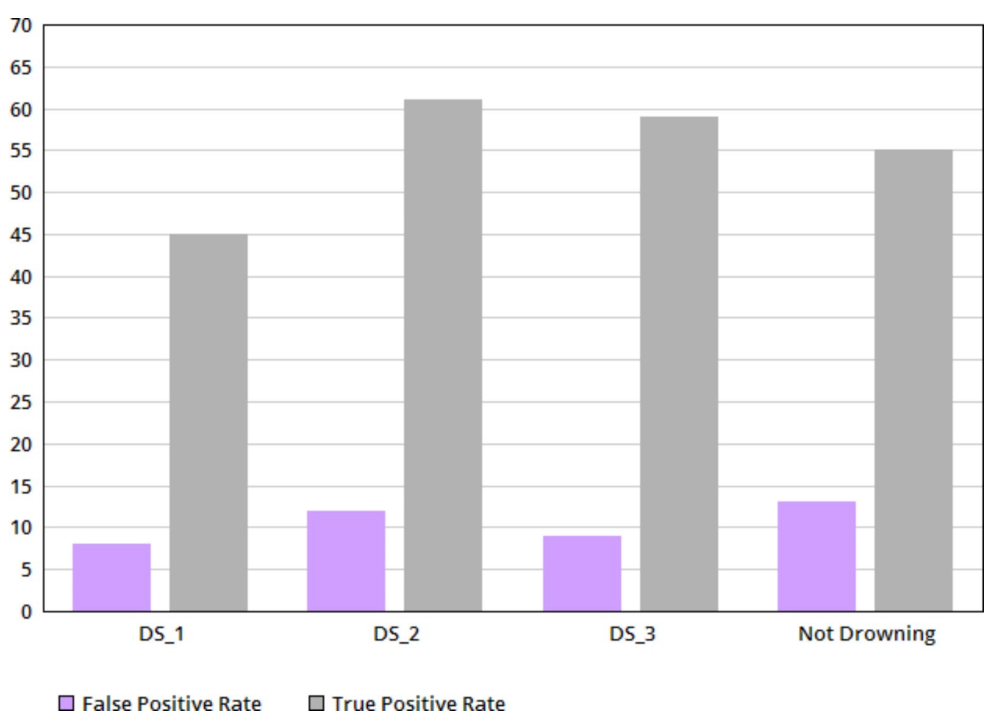

Fig. 4 False positive vs. true positive rates for different stages of classification of drowning

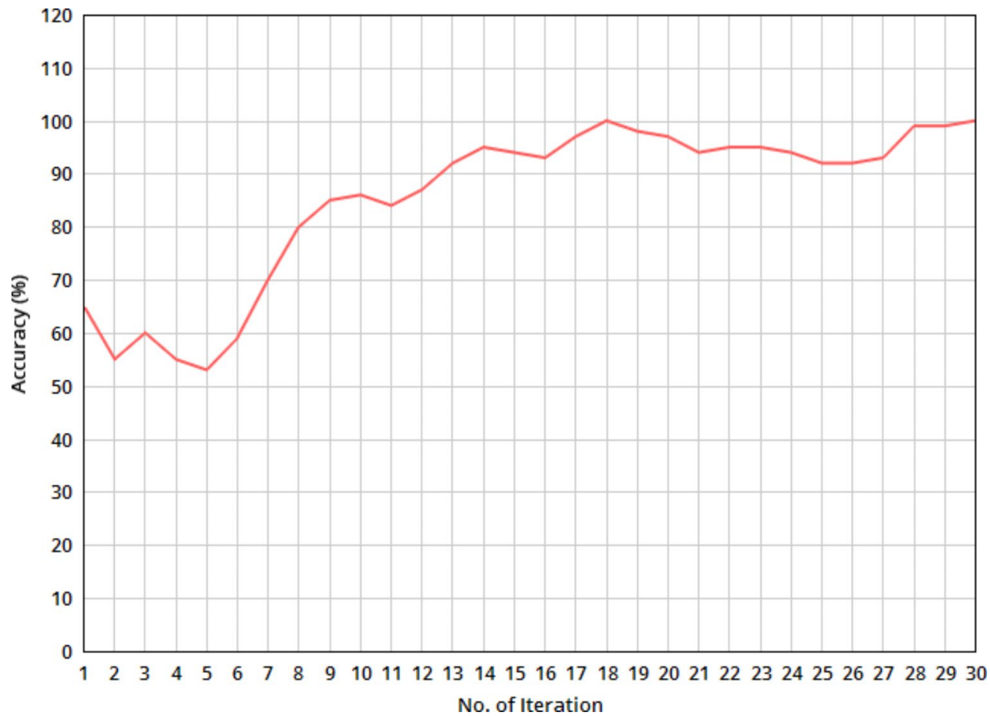

Fig. 5 Accuracy of the proposed methodology

the verification efficiency was 91.67%. The system retraining takes one minute as well as five seconds to accomplish, as well as a five-repetition approach will be used for verification.

4 Conclusion

The rising popularity of coastal destinations, combined with the inherent risks associated with water activities, necessitates innovative safety measures. Our research introduced an AI-Driven Drowned-Detection System that harnesses the power of deep learning algorithms to swiftly detect potential drowning incidents. The comprehensive device, which seamlessly amalgamates these elements, not only assists lifeguards in identifying drowning victims but also efficiently dispatches drones to deliver life-saving tubes and pinpoint perilous activities. The primary end-user, likely a pool owner or manager, will benefit from an integrated application, equipped with warning signals and a wireless system connected to Firebase notifications, ensuring prompt alert mechanisms. However, the system does have its limitations. The reliance on a single camera might restrict its field of view, potentially compromising detection quality. This drawback can be mitigated by deploying multiple cameras positioned strategically across diverse vantage points, enhancing the visual accuracy of the system. Another consideration is the operational constraint of drones under adverse weather conditions like heavy rain, gusty winds, or lightning. While swimming in such extreme conditions is generally discouraged, the system's adaptive capabilities are evident as it can proactively send out warnings when someone ventures into the water during such unfavorable scenarios, negating the need for aerial drone surveillance. As more individuals and families are drawn to the allure of the seashore, the integration of such technologically advanced systems becomes imperative not just for lifesaving interventions but also for pre-emptive safety alerts. This research epitomizes the strides being made in coastal safety, emphasizing the ongoing need for technological evolution and innovation in this critical domain.

Funding Authors declare that they have not received any funding for the present work.

Data availability Data included in article/supplementary material/referenced in article.

Declarations

Competing interests The authors have no competing interests to declare that are relevant to the content of this article.

References

1. WHO (2014). Global Report on Drowning: Preventing a Leading Killer; World Heath Orgnization: Geneva, Switzerland.
2. World Health Organization. World Drowning Prevention Day (2022). Available online: https://www.who.int/campaigns/worlddrowning-prevention-day/2022 (accessed on 20 April 2022).

3. Mohtasham-Amiri, Z. (2022). Traumatic injuries in drowning. J. Inj. Violence Res. 14, 6, PMCID: PMC9115828.
4. Zaara, M., Belhaj, A., Naceur, Y., Makni, C., Gharbaoui, M., Bellali, M., Zhioua, M., & Allouche, M. (2022). Patterns of unintentional fatal drowning among children in North Tunisia: A 10-year study. *Rev D'épidémiologie St Publique*, *70*, 31–37.
5. Alotaibi, A. (2020). Automated and Intelligent System for Monitoring Swimming Pool Safety based on the IoT and transfer learning. *Electronics*, *9*, 2082.
6. Li, D., Yu, L., Jin, W., Zhang, R., Feng, J., & Fu, N. (2021). An Improved Detection Method of Human Target at Sea Based on Yolov3. In Proceedings of the IEEE International Conference on Consumer Electronics and Computer Engineering, Guangzhou, China, 15–17 January.
7. Alshbatat, A. I. N. (2020). "Automated Vision-based Surveillance System to Detect Drowning Incidents in Swimming Pools.," in In Proceedings of the 2020 Advances in Science and Engineering Technology International Conferences (ASET), Dubai, UAE.
8. Prakash, B. D., Near-drowning early prediction technique using novel equations (NEPTUNE) for swimming pools, IAG Firemark Singapore.
9. Fares Jalled, V., & Voronkov, I. (2016). "Object detection using image processing," in arXiv preprint arXiv:1611.07791.
10. H. M. A., "The Swimmers Motion Detection Using Improved VIBE Algorithm," in International Conference on Robotics and Automation in Industry (ICRAI), (2019).
11. How-Lung, E., "An automatic drowning detection surveillance system for challenging outdoor pool environments".
12. Wenjie Zhan, Z. M. (2020). "From Body Parts Responses to Underwater Human Detection: A Deep Learning Approach,".
13. Carballo-Fazanes, A. (2020). The visible behaviour of drowning persons: A pilot observational study using Analytic Software and a nominal group technique, *International Journal of Environmental Research and Public Health p* 17.
14. Gopalakrishnan, A. (2019). "A neural temporal model for human motion prediction," in IEEE/CVF Conference on Computer Vision and Pattern Recognition.
15. Redmon, J. (2016). "You Only Look Once: Unified, Real-Time Object Detection," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
16. Jeon, E. I., Yu, S. J., Seok, H. W., et al. (2017). Comparative evaluation of commercial softwares in UAV imagery for cultural heritage recording: Case study for traditional building in South Korea. *Spat Inf Res*, *25*, 701–712. https://doi.org/10.1007/s41324-017-0137-z.
17. Yang, C., "Human-Pose-Estimation-Benchmarking-and-Action-Recognition," Github, 10 March 2019. [Online]. Available: https://github.com/ChengeYang/Human-Pose-EstimationBenchmarking-andActionRecognition/blob/master/doc/command_benchmarking.md. [Accessed 10 10 2021].
18. Sey, N. E. N., Amo-Boateng, M., Domfeh, M. K., et al. (2023). Deep learning-based framework for vegetation hazard monitoring near powerlines. *Spat Inf Res*. https://doi.org/10.1007/s41324-023-00518-0.
19. Hidalgo, G. (2019). "OpenPose: Whole-Body Pose Estimation,".
20. Google Play, "HelpStork," Google, 26 August 2021. [Online]. Available: https://play.google.com/store/apps/details?id=com.automatedlifeguard.push_notifs. [Accessed 12 10 2021].
21. Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genom*, *21*, 6.
22. Chicco, D., Starovoitov, V., & Jurman, G. (2021). The benefits of the Matthews correlation coefficient (MCC) over the diagnostic odds ratio (DOR) in binary classification assessment. *Ieee Access : Practical Innovations, Open Solutions*, *9*, 47112–47124.
23. Boughorbel, S., Jarray, F., & El-Anbari, M. (2017). Optimal classifier for imbalanced data using Matthews correlation coefficient metric. *Plos One*, *12*, e0177678.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.