Drowning Recognition for Ocean Surveillance using Computer Vision and Drone Control

Maureen Rakotondraibe, Tianyang Fang, and Jafar Saniie
Embedded Computing and Signal Processing (ECASP) Research Laboratory (http://ecasp.ece.iit.edu)
Department of Electrical and Computer Engineering
Illinois Institute of Technology, Chicago IL, U.S.A.

***Abstract*—According to WHO's report from 2021, Drowning is the 3rd leading cause of unintentional death worldwide. The use of autonomous drones for drowning recognition can increase the survival rate and help lifeguards and rescuers with their life saving mission. This paper presents a real-time drowning recognition model and algorithm for ocean surveillance that can be implemented on a drone. The presented model has been trained using two different approaches and has 88% accuracy. Compared to the contemporary models of drowning recognition designed for swimming pools, the model presented is better suited for outdoor applications in the ocean.**

Keywords—Drowning Recognition, Ocean Surveillance, Computer Vision, Artificial Intelligence, Deep Learning

I. Introduction

Autonomous drones are unmanned aerial vehicles (UAVs) that operate using Artificial Intelligence (AI) and do not require a human pilot. Over the past years, the market of autonomous drones has expanded a lot. Recently, we can observe breakthroughs in the development and implementation of drone deliveries around the world. Amazon has scheduled a test phase in California by the end of the year for its new delivery service, Amazon Prime Air [1]. Wing by Alphabet achieved some flights in Texas, Finland, and Australia [2]. Finally, Skyports is drone-delivering with RoyalMail in the UK and facilitating ship-to-shore transportation from cargo in Singapore [3]. Autonomous drones are also used for industrial applications. The top 5 industries using them are mining, solar energy, oil and gas, thermoelectric power, and ports and terminals. Drones are convenient for site inspections, data collection, and anomaly detection [4]. For example, drones can be used for gas and fire inspection in a big infrastructure with inaccessible areas.

Apart from deliveries and industrial uses, autonomous drones can be used for search and rescue missions. They can be used for avalanche rescuing, earthquakes, collapsed buildings, firefighters, and rescue at sea. Using drones in those missions can help reduce search time, increase rescuers' visibility, and increase the chance of survival. This research focuses on rescue at sea missions. Drowning is a deadly and costly issue worldwide:

- There are around 236,000 people that die from drowning in the world each year. It is the 3rd leading cause of unintentional injury death. [5]
- 4 coastal deaths every 5 days between December 2021 and February 2022 in Australia [6]
- 44% of fatal drownings in France occurred by sea in 2020 [7]
- Coastal drowning in the United States accounts for \$273 million yearly in direct and indirect costs. In Australia and Canada, the total annual cost of drowning injuries is \$85.5 million and \$173 million, respectively. [5]

Drowning can happen fast. The longer it takes to detect and save a person, the higher the risk of death. The motivation of this project is to increase the survival rate of drowning people in the ocean. To achieve this, a drone is used for live video feed over the ocean area, and computer vision is used to recognize and locate people at risk and provide floating aids. This paper presents a drowning recognition model suited for an autonomous drone solution to decrease the response time and help the rescuers have more visibility.

A. Contemporary Work for Drowning Detection

Drowning detection has been a popular topic in the past two decades. Early researchers utilize machine vision techniques to analyze photos of the pool to achieve the goal of drowning detection [8][9][10][11]. In reference [8], the drowning detection algorithm recognizes struggles on the water surface to identify early drowning behavioral signs from video clips, where hidden Markov Models (HMM) are used as one of two methods to recognize drowning behavioral signs. The paper claimed to achieve promising results, although this algorithm is subject-dependent and used a small dataset. In reference [9][10][11], the surveillance system uses the posture, orientation, speed, and movement range of the swimmers to detect drowning behavior. After swimmers are extracted from the aquatic environment with image processing techniques, a decision model based on specific descriptors inspired by lifeguards' feedback determines if a drowning event exists. These papers also presented promising results with models more resilient to crowded scenarios, however, they need good visibility and clear water due to the silhouette extraction algorithms used to identify and track each swimmer. While being efficient for swimming pools, those approaches are not suited for rescue in the ocean, where the water has much movement and a darker color due to the depth and where the area to monitor is significantly bigger. In recent years, deep learning models have been used to get better performances. In reference [12], the system uses Thin-MobileNetv2 (CNN) for human recognition and OpenPose for skeleton recognition to identify drowning scenarios using posture recognition. This system achieved an accuracy of 89.4%. However, the model was tested using only one swimmer, which makes the system subject-dependent. Furthermore, the camera was placed underwater at the bottom of a swimming pool which limits the application of this system for ocean drowning recognition.

This kind of system's drawbacks are high installation costs and limited to a closed environment. The presented system uses deep learning where cameras are onboard drones to extend the range search and address deep water drowning scenarios without installation constraints.

B. *Drowning Recognition Model*

This research aims to design a drowning recognition model for ocean surveillance and rescue in the event of people in danger in the sea. It will help the lifeguards locate the people in danger in the water and help the victims by dispatching a lifeline in order to help them wait for the rescuers' team.

Fig. 1 shows the system flowchart of the drowning surveillance system. The embedded camera in the drone will be used to monitor a specific area. The video stream will be analyzed by the drowning recognition model, which will be able to identify and locate the people in danger. If a person in danger is recognized, the drone will be instructed to dispatch a lifeline next to the person and always keep the victim in sight to report the person's localization to the rescuers.

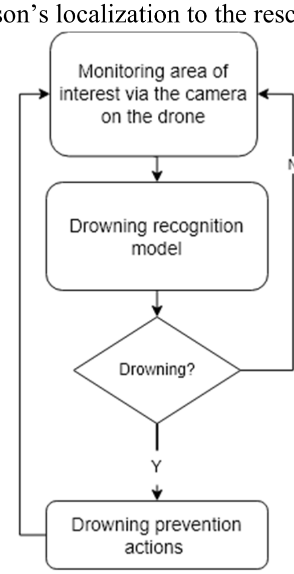

Fig. 1. Drowning Surveillance System Flowchart

In the following paragraphs, the system design will be presented, including a drone study for this application, the main system algorithm, and the drowning recognition model design.

II. Drowning Recognition System Design

A. *Drone Control*

The drowning recognition model can be implemented either on a base station or directly onboard the drone if it has the necessary features. Two different drones have been studied for this application. The first drone option is the ANAFI Parrot drone. Table I is a summary of the key specifications.

TABLE I. SPECIFICATIONS OF THE ANAFI PARROT DRONE [13]

Visible Imaging System	**Flight Envelope**	**Flight Modes**
RGB sensor: CMOS 1/2,4''	Flight time: 26 min	Visual tracking (Cameraman)
Max resolution: 4K HDR	Max vertical speed: 4m/s	Point of interest tracking (Touch & Fly)
Stabilization: 3-axis hybrid	Max speed: 55 km/h	Pilot tracking (Follow Me): In-app purchase
Zoom: x 3 (lossless: x 2,8)	Wind resistance: 50km/h	Flight programming (Flight Plan): In-app purchase

The 4K camera and the zoom abilities will capture high-quality images important for human recognition. Wind resistance is important for rescue at sea missions since wind can be present in coastal areas or the middle of the sea. The flight time should be sufficient since it takes 5 to 6 minutes on average for people to drown. This time should be enough for locating the people and giving them floating devices so that the rescuers have more than 5-6 minutes to intervene. Finally, the point of interest tracking is useful since the system should be able to track the victims in the water.

The second option is the ANAFI AI drone. Compared to the ANAFI 4K, the main advantages are that it can run the AI algorithm on board and has a better zoom (x6), 32 minutes of battery life, and 4G connectivity.

B. *System Algorithm*

The system algorithm is the same whether the drowning recognition model runs onboard the drone or on a base station. The drone will capture a video stream and will transmit it to a video analysis module. The video analysis module will use the drowning recognition model based on the YOLOv5 architecture to identify and locate people at risk. and will send the results to a drone control module. If people at risk are detected, the drone will hover above the person to release a floating aid and track the localization. Fig. 2 presents the workflow of the main system algorithm and the communication between the different modules.

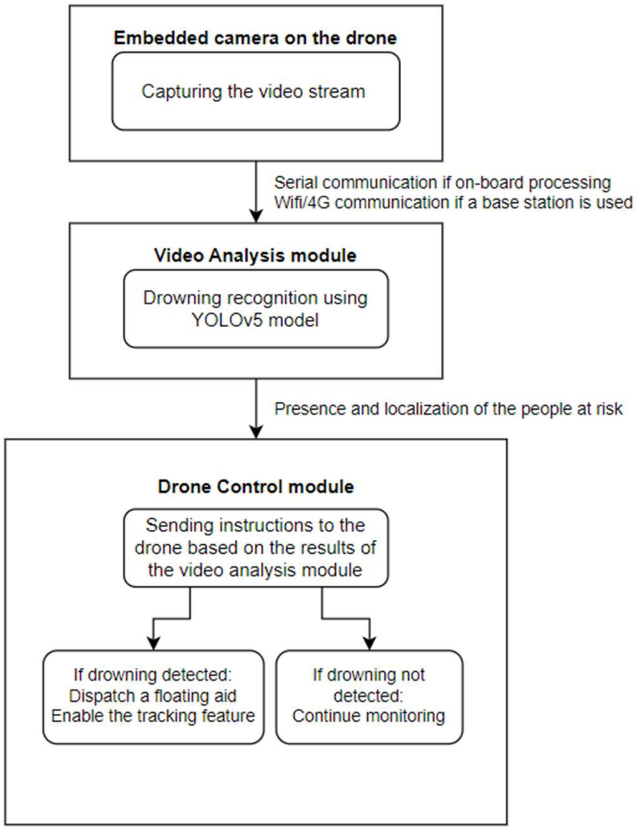

Fig. 2. Workflow of the main system algorithm and the communication between the different modules

C. Drowning Recognition Model Design using YOLOv5

Due to currents and waves, it is very difficult to spot people when they are swimming in the middle of deep water even for professional lifeguards [14]. Even if they are not drowning, being in that position makes them extremely vulnerable. Rescuing them quickly before they drift too far is a key objective. To achieve this objective, the speed of the algorithm and accuracy are critical.

To address the situation, the model trained is using the YOLOv5 architecture. YOLOv5 is an object detection model released in June 2020 by Ultralytics and is an evolution of previous YOLO architectures [15]. The name stands for 'You only look once' and is related to its single-stage architecture. It is popular for its optimized trade-off between speed and accuracy. In [15], it has also been identified as the most accurate YOLO architecture for UAV object detection in UAV images.

The head of the network is composed of four convolution layers that output the bounding boxes, the confidence scores, and the object classes. To improve the performance, the model uses two different loss functions. Loss functions are a critical part because it evaluates the model predictions during training and allows the adjustments of the neural network weights. The Binary Cross Entropy function computes the loss for class loss and the objectness loss. Objectness refers to the probability that an object exists in a region of interest. The second loss function uses Complete Intersection over Union to compute the location loss. It correlates overlapping boxes to class prediction for the most precise bounding box.

To train the YOLOv5 model on a custom dataset, the following parameters are needed.

- Dataset: the dataset must be labeled and split into three sets. The training set is used for learning purposes. The validation set is used to tune the performance of the model. The testing set is used to get the actual performance of the model.
- Preloaded parameters (optional): pre-trained weights can be used to speed up the training process
- Epochs: number of training iterations
- Batch size: number of samples processed in one iteration
- Learning rate: how much the weights in the model should change to minimize the error

The workflow of the drowning recognition model is shown in Fig. 3. The first stage consists of the labeled dataset and the data pre-processing; the second stage involves the training and the validation; finally, in the last stage a performance evaluation is carried out.

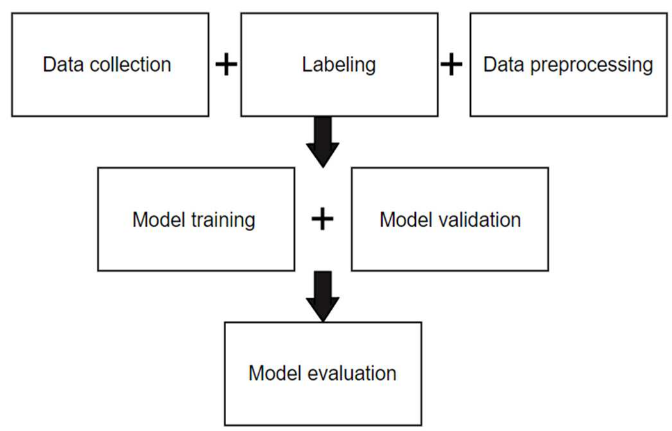

Fig. 3. Workflow of the Drowning Recognition Model Design

D. Model Training

To train a drowning recognition model, a custom dataset is needed. Data collection and labeling were done using Roboflow. Roboflow is a computer vision developer framework designed to help with data collection, preprocessing, and model training techniques [16].

The first approach was to choose 100 photos that applied to the drowning in the ocean scenario. Fig. 4 is an example of how humans in deep water are labeled at risk. They are highlighted by yellow boxes.

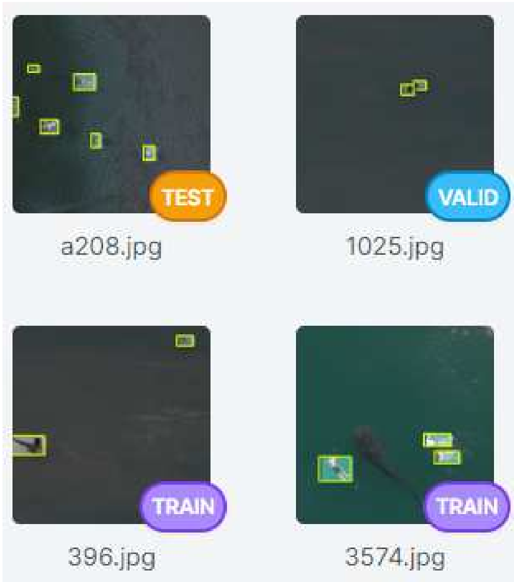

Fig. 4. Example of Labeled Images in the Dataset

Roboflow also makes it easier to split data into training, validation, and testing sets. Fig. 5 is an example of the data repartition within the Roboflow application interface, where 70% of images were put in the training, 20% in the validation, and 10% in the testing set.

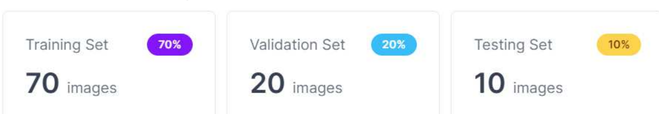

Fig. 5. Example of Data Repartition using Roboflow Framework

The images are all resized to 640x640 pixels for faster training and better performance. To improve the resilience of the model to camera focus and moving targets, data augmentation is used. It allows the creation of additional data by altering some features. A Gaussian blur of up to 10 pixels was added to create additional images.

After the data collection step and the data preprocessing steps, parameter tuning is used to improve the performance. It is called the validation step. In this step, I modified the batch size, the learning rate, and the number of epochs. To accelerate the training process, I also used pre-trained weights. I used the weights of a model trained to recognize 80 categories from the COCO dataset. After refining the parameters, I obtained 75% of accuracy after 100 epochs, with 100 samples and parameters tuning. This process is challenging and needs many iterations. The evolution of the accuracy for some of the trials is shown in Fig. 6. For example, the model reached overfitting in the blue trial because the data augmentation via pre-processing was too big compared to the comparably small database. Overfitting means that the model memorizes the features instead of learning how to recognize them, leading to poor performance when confronted with new data.

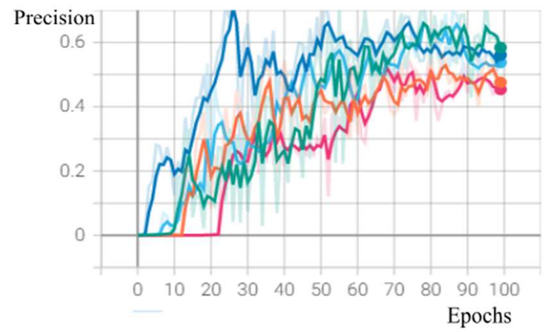

Fig. 6. Evolution of the Precision through the Number of Epochs during Parameters Tuning

The second approach to training this model is to use a bigger database. With 1859 samples, the precision obtained was 88% after 100 epochs without the need to tune the parameters. These two approaches show that to train a model for a specific application if the dataset available is small, it is possible to obtain good accuracy, but it takes time and parameter tuning. The bigger the database, the faster the process.

III. Tested Results

After the validation process, the model was tested on new data to evaluate the actual performance. The results are shown in Table II.

TABLE II. RESULT OF THE MODEL WHEN TESTED WITH 10 DIFFERENT IMAGES CONTAINING 30 PEOPLE IN TOTAL IN DIFFERENT DROWNING POSITIONS

Image Index	The person at risk recognized by the model	Ground Truth	Error
1	1	1	0
2	6	6	0
3	2	2	0
4	3	3	0
5	3	2	+1
6	6	4	+2
7	4	4	0
8	3	3	0
9	2	2	0
10	3	3	0

The model has been tested on 10 images containing 30 people in the water overall. The model perfectly recognized 24/30 people. The model recognizes three more people by error, but all 30 people have been recognized and located (80% of perfect accuracy, 100% people at risk detected). An example image of the output result is shown in Fig. 7.

Fig. 7. Example of the Output of the Drowning Recognition Model

IV. CONCLUSION AND FUTURE WORK

This research is a rescue-at-sea autonomous system. Despite the work of rescuers at sea and lifeguards in coastal areas, drowning continues to be deadly and costly and affects people worldwide. The presented paper is innovative because it uses drones and computer vision to help the rescuers spot people in the sea and give them more time to save everyone by dropping a life vest or floating device and tracking the victims. In the market, drowning detection devices are only suitable for swimming pools and search and rescue drones usually use either thermal imagery or phone localization to detect people. Those technologies are not suitable for people in the water.

The human in deep sea model implemented is a YOLOv5 custom-trained model. The challenge was identifying humans in unclear water from a drone view. Recognizing people in the water is difficult because of the moving waves, the zoom and camera focus, and the body parts immersed. Two approaches were studied. In the first one, a small database and parameter tuning was used. This model reached 75% of accuracy. In the second one, a bigger database of 1859 samples were used, and the accuracy reached 88% without parameters tuning. To implement this model, different drones are proposed. The ANAFI 4K is suitable but discontinued. The model can run on a base station and communicate with the drone via WI-FI. However, the more suitable and recommended choice is the ANAFI AI since it has better zoom, has 4G connectivity and can run the AI model onboard.

For future work, it is recommended to add a behavioral analysis to the model and a condition analysis to prioritize help in the case several people are in the water.

REFERENCES

[1] A. Staff, “Amazon prime air prepares for drone deliveries,” *US About Amazon*, 13-Jun-2022. [Online]. Available: https://www.aboutamazon.com/news/transportation/amazon-prime-air-prepares-for-drone-deliveries. [Accessed: 2022].

[2] A. J. Hawkins, “Alphabet's wing kicks off drone delivery service in Dallas on April 7th,” *The Verge*, 04-Apr-2022. [Online]. Available: https://www.theverge.com/2022/4/4/23006894/alphabet-wing-drone-delivery-dallas-kick-off. [Accessed: 2022].

[3] B. Crumley, “Skyports in cargo drone delivery infrastructure deal with Singapore's Jurong Port,” *DroneDJ*, 24-May-2022. [Online]. Available: https://dronedj.com/2022/05/24/skyports-in-cargo-drone-delivery-infrastructure-deal-with-singapores-jurong-port/. [Accessed: 2022].

[4] “The ultimate guide to autonomous drones for industrial sites,” Percepto, 24-Aug-2022. [Online]. Available: https://percepto.co/autonomous-drone/. [Accessed: 2022].

[5] “Drowning,” World Health Organization, 27-Apr-2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/drowning. [Accessed: 2022].

[6] “SLSA Summer Coastal Drowning Report: Four coastal deaths every five days of Summer,” Surf Life Saving Australia, 18-Mar-2022. [Online]. Available: https://sls.com.au/slsa-summer-coastal-drowning-report-four-coastal-deaths-every-five-days-of-summer/. [Accessed: 2022].

[7] “Quatrièmes Résultats de l'enquête NOYADES 2018 en France,” Accueil, 05-Sep-2018. [Online]. Available: https://www.santepubliquefrance.fr/les-actualites/2018/quatriemes-resultats-de-l-enquete-noyades-2018-en-france. [Accessed: 2022].

[8] Wenmiao Lu and Yap-Peng Tan, "Swimmer motion analysis with application to drowning detection," 2002 IEEE International Symposium on Circuits and Systems (ISCAS), Phoenix-Scottsdale, AZ, USA, 2002, pp. II-II, doi: 10.1109/ISCAS.2002.1011439.

[9] Eng, Toh, Kam, Wang and Yau, "An automatic drowning detection surveillance system for challenging outdoor pool environments," Proceedings Ninth IEEE International Conference on Computer Vision, Nice, France, 2003, pp. 532-539 vol.1, doi: 10.1109/ICCV.2003.1238393.

[10] Wenmiao Lu and Yap-Peng Tan, "A vision-based approach to early detection of drowning incidents in swimming pools," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 14, no. 2, pp. 159-178, Feb. 2004, doi: 10.1109/TCSVT.2003.821980.

[11] H. -L. Eng, K. -A. Toh, W. -Y. Yau and J. Wang, "DEWS: A Live Visual Surveillance System for Early Drowning Detection at Pool," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 18, no. 2, pp. 196-210, Feb. 2008, doi: 10.1109/TCSVT.2007.913960.

[12] J. -X. Jian and C. -M. Wang, "Deep Learning Used to Recognition Swimmers Drowning," 2021 IEEE/ACIS 22nd International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), Taichung, Taiwan, 2021, pp. 111-114, doi: 10.1109/SNPD51163.2021.9704884.

[13] “Anafi Parrot Drone - all of the technical specifications,” Parrot. [Online]. Available: https://www.parrot.com/us/drones /anafi/technical-specifications. [Accessed: 2022].

[14] A. Roy and K. Srinivasan, "A novel drowning detection method for safety of Swimmers," 2018 20th National Power Systems Conference (NPSC), Tiruchirappalli, India, 2018, pp. 1-6, doi: 10.1109/NPSC.2018.8771844.

[15] T. S. Gunawan, I. M. M. Ismail, M. Kartiwi and N. Ismail, "Performance Comparison of Various YOLO Architectures on Object Detection of UAV Images," 2022 IEEE 8th International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA), Melaka, Malaysia, 2022, pp. 257-261, doi: 10.1109/ICSIMA55652.2022.9928938

[16] “Roboflow Overview,” Roboflow. [Online]. Available: https://docs.roboflow.com/. [Accessed: 2022].