The Use of Synthetic Media in Autonomous Aerial Search and Rescue

Minh Nguyen and Dr. Jonathan Adams, Odirile Moja, Christopher Ko
College Of Communication And Information, Florida State University, Tallahassee, Fl 32303

Abstract

According to the World Health Organization, drowning claims approximately 300,000 lives globally each year. Unmanned aerial vehicles (UAVs) integrated with machine learning (ML) have shown promise in drowning victim search-and-rescue efforts. However, ML models depend heavily on diverse and annotated datasets, which are challenging to obtain due to financial and ethical constraints. This study evaluates the effectiveness of synthetic visual data, generated from 3D models and backgrounds, for training a YOLOv9 object detection model. The synthetic dataset, combined with real images, will be tested and compared to results from “The use of drones and a machine-learning model for recognition of simulated drowning victims—A feasibility study.” Model performance will be assessed based on recognition rates, detection accuracy, and combined ML-human identification accuracy.

Introduction

In recent years, the adoption of unmanned aerial vehicles (UAVs) equipped with live improvements alone cannot fully address these acvideo has expanded rapidly, revolutionizing up to 15 industries (Drone applications, 2024). In maritime operations, UAVs significantly enhance capabilities for monitoring large open-water areas (Varga et al., 2022). Integrating machine learning (ML), particularly convolutional neural networks (CNNs), with UAVs has shown effectiveness for critical tasks (Claesson et al., 2020). However, CNN performance is limited by dataset scarcity and lack of object diversity, as even extensive datasets may not adequately represent edge cases, and computational curacy gaps (Adams et al., 2021). Collecting large, diverse datasets is costly, time-consuming, and prone to annotation errors, with high-quality data often particularly limited (Talukdar et al., 2018). Synthetic media thus emerges as a promising approach, offering diverse, high-quality, customizable datasets to supplement real-world data. This research evaluates the effectiveness of synthetic media as an alternative to costly authentic data in ML model training.

Methodology

The research is structured into three key phases. In the first phase, we implement a 3D model generation pipeline that synthesizes realistic images by combining 3D human models in object file format with various maritime background images.

Phase I: Produce high-fidelity images with diverse camera angles, positions, and environmental conditions that mimic authentic SAR scenarios.

Phase II: Synthetic images serve as training data for a YOLO object detection model. Three to eight percent are authentic images.

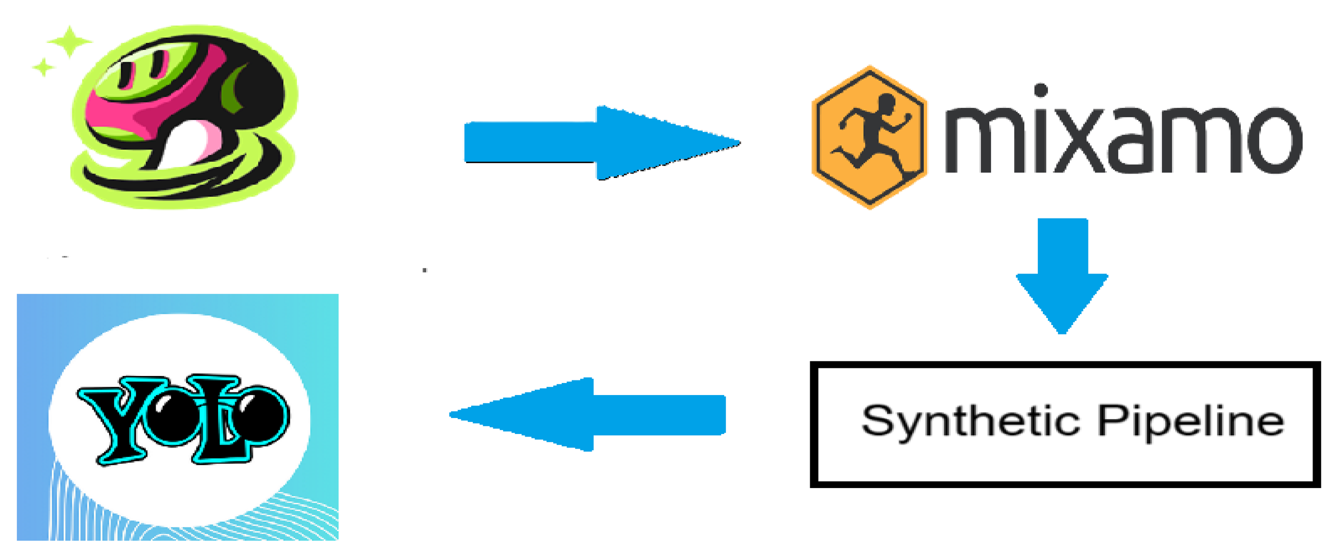
Rationale: The fusion of synthetic and real data ensures that the model accurately detects drowning victims across varied situations.

Evaluation: Comparative analysis with previous research on the recognition of simulated drowning victims, by the proportion of cases where the model successfully detects objects, and the overall accuracy when combining automated detection with manual verification.

To the left are the synthetically generated images by our pipeline, combining random backgrounds and human objects from our datasets. To the right is images taken from the study “The use of drones and a machine-learning model for recognition of simulated drowning victims—A feasibility study”

Results

While no definitive findings are expected yet, our research contributes to improved rescue outcomes and enhanced safety protocols in maritime environments, while paving the way for the development of innovative AI-driven emergency response solutions. This project not only streamlines the generation of training data but also provides a scalable, flexible framework for advancing rescue technology. By integrating advanced synthetic techniques with real-world data, our approach significantly reduces cost and logistical challenges typically associated with traditional SAR training datasets.



1) 3D models of drowning individuals are generated using Meshy. 2) Next, Maximo adjusts each model’s pose to accurately represent swimming or drowning postures. 3) Our synthetic pipeline then mass-produces visual datasets by combining these 3D models with diverse backgrounds. 4) Finally, YOLO is used to train and evaluate the detection model.

Resources

Claesson, A., et al., (2020a). The use of drones and a machine-learning model for recognition of simulated drowning victims—a feasibility study. Resuscitation, 156, 196–201.

Drone applications. Drone Industry Insights. (2024, October 9). https://droneii.com/drone-applications

Varga, L. A., Kiefer, B., Messmer, M., & Zell, A. (2022). Sea DronesSee: A maritime benchmark for detecting humans in Open water. *2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 3686–3696.

Adams, J., Sutor, J., Dodd, A., & Murphy, E. (2021). Evaluating the performance of synthetic visual data for real-time object detection. *2021 6th International Conference on Communication, Image and Signal Processing (CCISP)*, 167– 171.

Talukdar, J., Biswas, A., & Gupta, S. (2018). Data augmentation on synthetic images for transfer learning using DEEP CNNS. *2018 5th International Conference on Signal Processing and Integrated Networks (SPIN)*, 215–219.