StormVision

GenAI-Based Person-in-Water Detection in Rough Seas

Presenter:
Tomer Atia, Shalev Cohen, Noam Hadad

Motivation and Use Case

This project tackles the challenge of detecting people in the water during rough sea conditions, where visibility is low and failures are most likely.

PROJECT GOALS

- Generate synthetic rough-sea images to fill the gap in real storm data.
- Test whether this synthetic data can improve detection in challenging conditions.

LIMITATIONS OF CURRENT DATASETS

- Most datasets contain only calm-sea images.
- Rough-sea scenes are missing due to the danger of filming real storms.

NOVELTY

- No real rough-sea images exist due to the danger of filming storms.
- Create synthetic storm images with GenAI to fill this gap.
- Evaluate whether this improves detection in extreme conditions.

ML Task Definition and Objectives

The project performs a binary classification task to detect whether a person is present in the water using aerial drone images.

INPUT& OUTPUT

- **Input**: RGB drone image of sea
- **Output**: Binary label:
  - 0 = no person
  - 1 = person in water

Classification Model

EfficientNet-BO (pretrained)

- Lightweight, easy to fine-tune
- Good for small datasets and small objects
- Ideal for aerial drone imagery

Training Method

- Replace head → 2 classes
- Partial fine-tuning of top layers

Dataset & Synthetic Data

DATASET

- We use a subset of a public maritime drone dataset (e.g., SeaDronesSee).
- Contains calm-sea images with and without people.
- This real data serves as the base for generating synthetic rough-sea samples.

DATA SPLITS

- **Real data** (calm sea):
  60% train_real / 20% val_real / 20% test_real
- **Synthetic storm data** (for training):
  80% train_synth / 20% val_synth
- **Synthetic Storm Data - Test (Separate Set):**
  100% test_storm_synth
  (Not used for training - evaluation only)

GENAI PIPELINE

- **Inpainting** - keeps the real person, replaces the sea with a storm version.
- **Full Generation** - creates fully synthetic rough-sea scenes.

Together they provide storm-condition images missing from real data.

Evaluation and Metrics for Model Performance

MODEL COMPARISON

- **Model A (Baseline model):** trained only on real calm-sea data
- **Model B (Storm model):** trained on real + synthetic rough-sea data

TESTING FOCUS

Models are evaluated on two test sets:

- **Test_real** (real calm-sea images)
- **Test_storm_synth** (synthetic rough-sea images)

This shows how synthetic storm data affects performance.

EVALUATION METRICS

Key metrics:

- **Accuracy -** overall correctness
- **Recall -** how many real “person in water” cases were detected
- **Precision -** how many detections were correct
- **F1 score -** balance between precision and recall