Towards a floating plastic waste early warning system

Gábor Paller and Gábor Élő
Széchenyi István University
INFORMATION SOCIETY RESEARCH & EDUCATION GROUP (ISREG)

Plastic in rivers

- 50-70% of all the solid pollution in rivers is plastic
- Size varies from microplastic (< 0.1 μm) to macroplastic (>5 cm)
- Relevant plastic pollution depends on the polluting entities
- In the rivers of East Hungary it is mostly mass macroplastic pollution from upstream waste dumps

Source: In the name of the Tisza - source to the Black Sea, https://www.youtube.com/watch?v=TLyK_aIu3fc

Response: heavy machinery

Environment

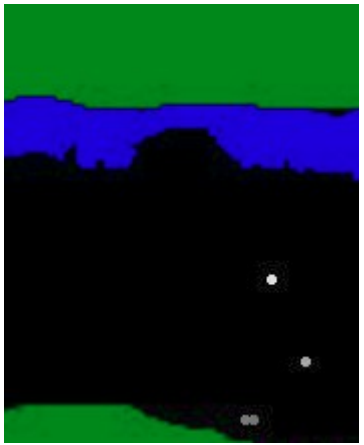
- Long detection distance (up to 50 meters)
- Plastic items covered with foreign materials (algae, water film, dirt, etc.)

- Observation in the visible light domain was chosen with infrared illumination option during the night

First iteration: motion detection

- Almost all the surveillance cameras have this option
- In order to try out all the possible configuration options, we built our own motion detection surveillance camera
  - Based on *motion* open-source software and Raspberry Pi 3
  - Supported by its own server backend deployed into Azure cloud
- Operated for more than a year, took more than 440000 images

Limitations

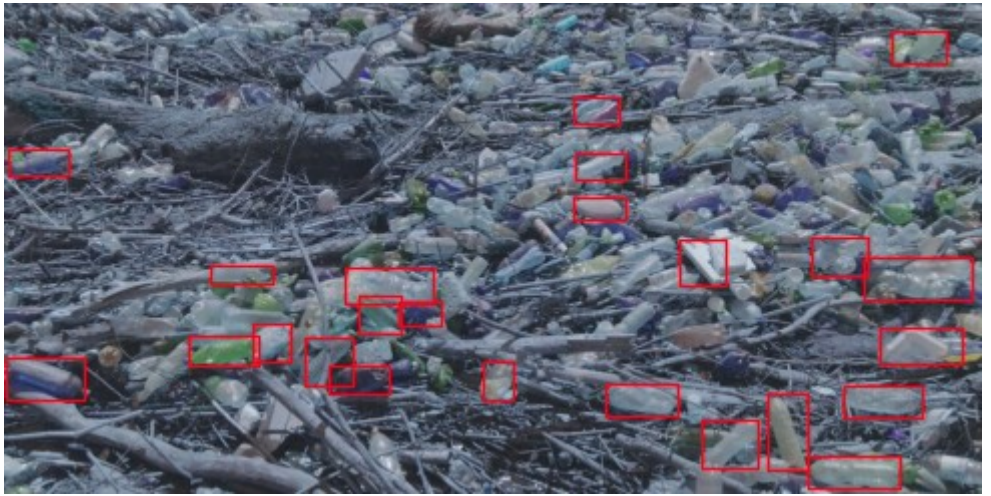


First iteration experiences

- Selectivity is very low, individual plastic waste items cannot be reliably recognized
- **Lot** of noise, most of the images are irrelevant
- Quite a good start for picking training images

Second iteration: deep learning

- Experimented with YOLOv3 and Faster R-CNN
- Finally chose Faster R-CNN with ResNet-50-FPN backbone
  - Needs less memory than ResNet-101 and it turned out to be crucial at training
- Image augmentation: rotating by 90,180 and 270 degrees.
- Zooming augmentation was not necessary due to FPN in the backbone



Trained with annotated images

Initial training results

ID	TR	TNR	TM	NRT	MR
#1	608	817	12	57.3%	1.97%
#2	1057	603	153	36.3%	14.47%
#3	309	129	31	29.4%	10.03%
#4	277	107	11	27.86%	3.97%
#5	0	0	0	N/A	N/A

TR - recognized
TNR - not recognized
TM - miscategorized
NRT - not rec. to all
MR - miscat. to rec.

The system

- Professional surveillance camera with 18x zoom
- GPU-equipped industrial PC in the camera unit
- Backend extended with retraining server
  - Human operator selects images where something relevant was not recognized and annotates them by hand
  - Retraining with new data added, new weights are sent to the cameras

Conclusions

- Second iteration significantly lowered the proportion of false alarms
- Due to large water surfaces, target objects often have bad quality images - perfect recognition is impossible
- False negatives (and some false positives) are still present
- Suitable for early warning system, not suitable for exact counting

Questions?