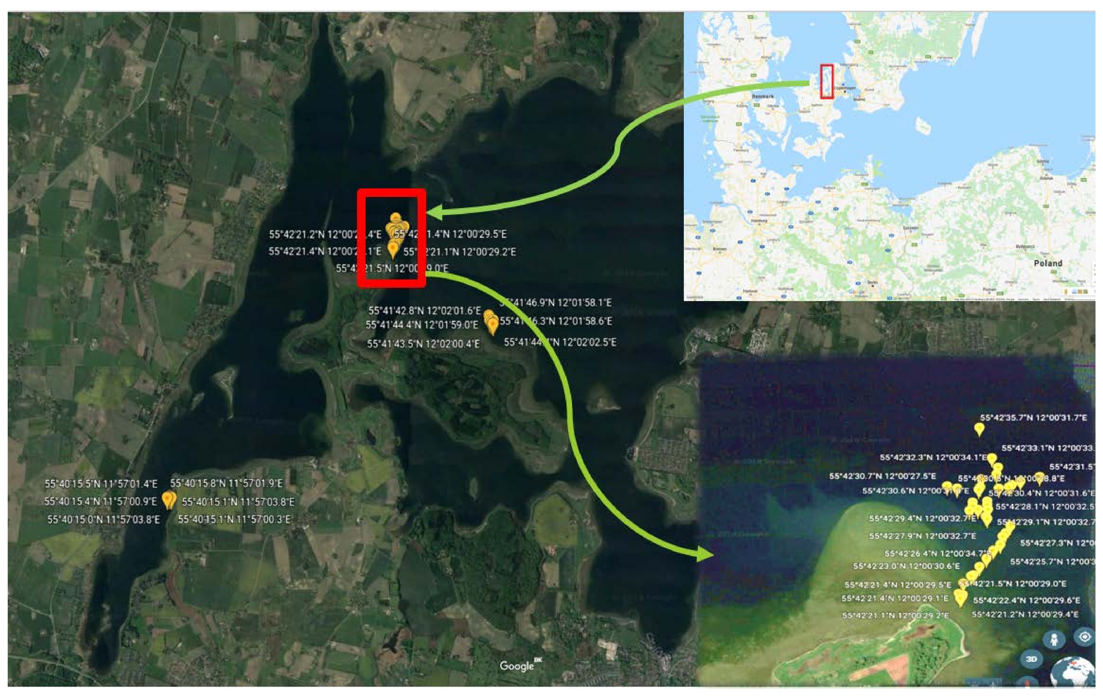
DTU Compute
Department of Applied Mathematics and Computer Science

UNIVERSITY OF COPENHAGEN

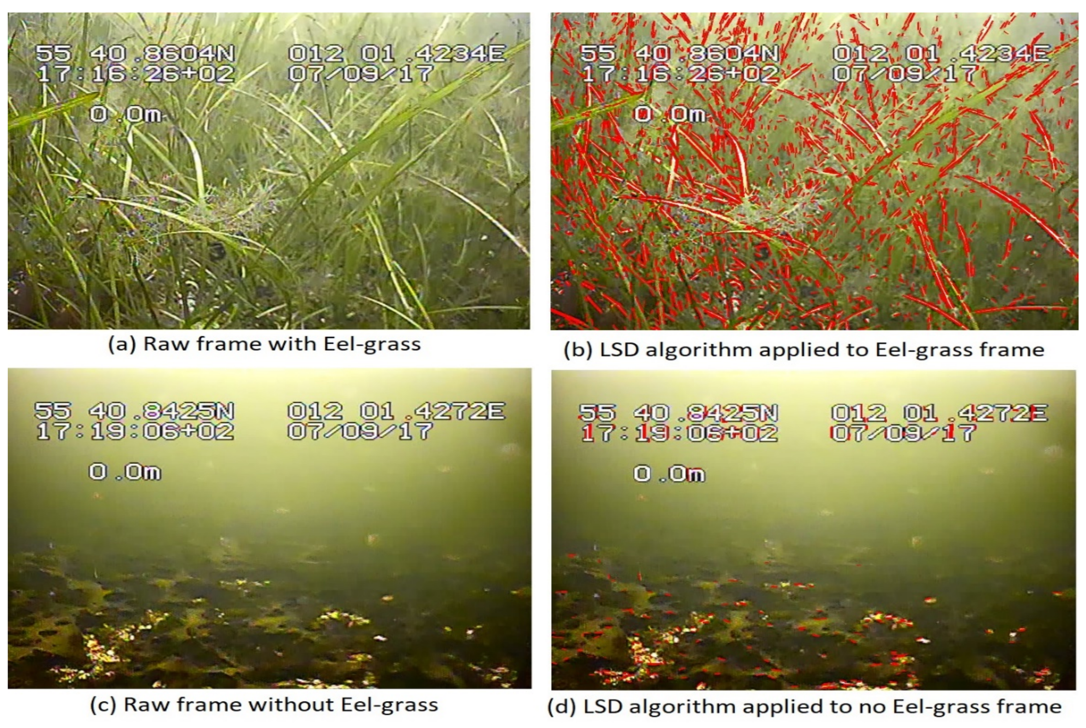
EelDetect: Unsupervised detection and quantification of eelgrass from under water videos

Sayantan Sengupta, Phd student, Statistics and Data Analysis
Bjarne K Ersbøll, Professor and Head of section, Statistics and Data Analysis
Anders Stockmarr, Associate Professor, Statistics and Data Analysis

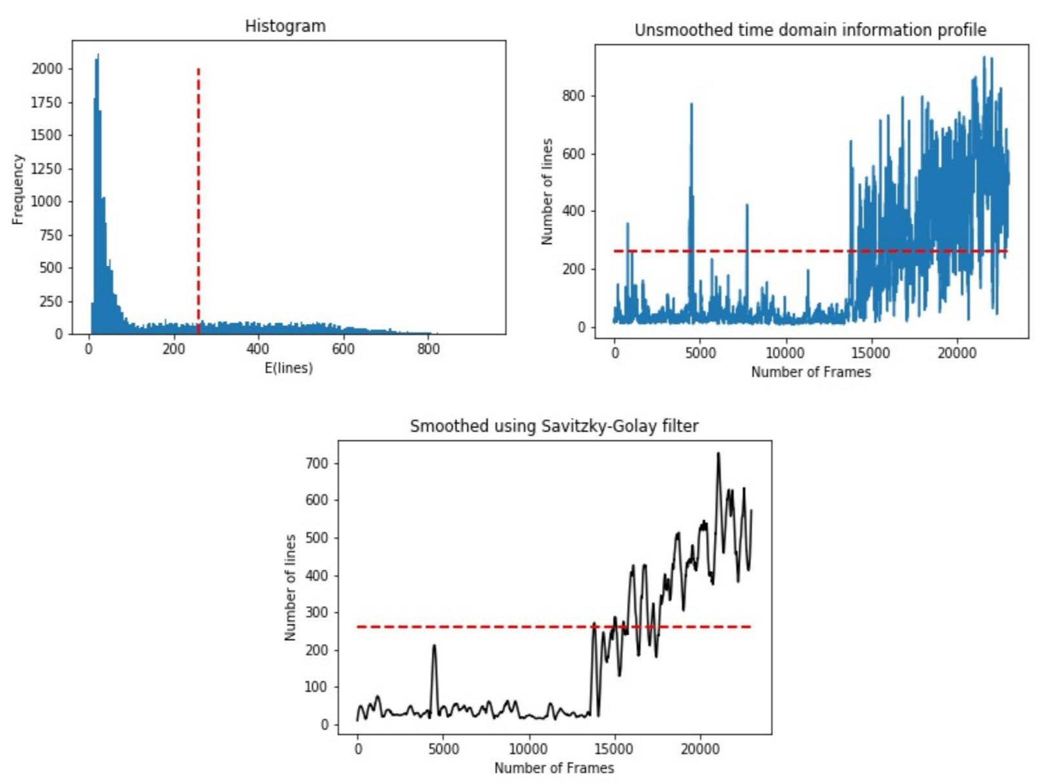
Overview: Benthic vegetation, an important indicator of the state of marine environment is difficult, expensive, error prone and subjective to monitor and quantify when done manually by domain experts. With advancement of ML techniques it is possible to automate the process. We use a line segment detector to extract features from under water videos and use two approaches to cluster into either eelgrass or no-eelgrass frames. We also extend this method to quantify the percentage coverage in a region by counting the ratio of present frames to total frames. The method is robust and is shown to rectify rare labelling mistakes from domain experts.



Map of the study site in Roskilde Fjord. Some of the marked locations where the data is available from under water videos.



Effect of LSD algorithm on two scenarios. Frames with eelgrass detects more number of Lines than frame without eelgrass.



The effect of smoothing on noisy fluctuations. Histogram of a video transect with number of lines detected. The red dashed line represents the threshold for distinguishing eelgrass and no eelgrass.

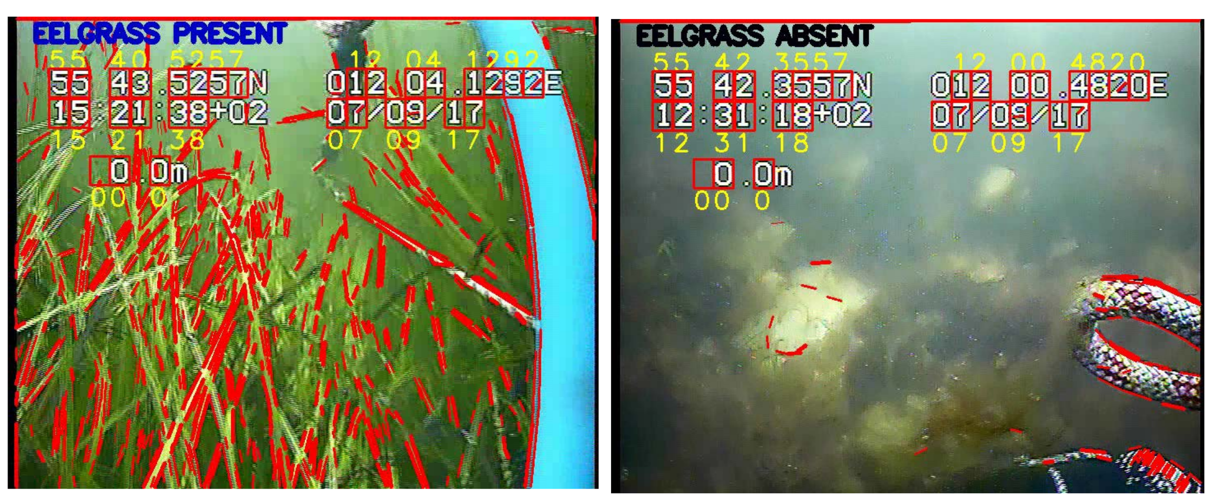
Heuristic approach:

- Extract two features: *Number of lines* and *Total length of lines* form each frames.
- Apply suitable smoothing (Savitzky-Golay filter).
- Find optimal threshold for *Number of lines* using visual inspection.
- Using the above find optimal *Total length of lines* which gives minimum mismatch of prediction.
- Find the optimal combination of the form,

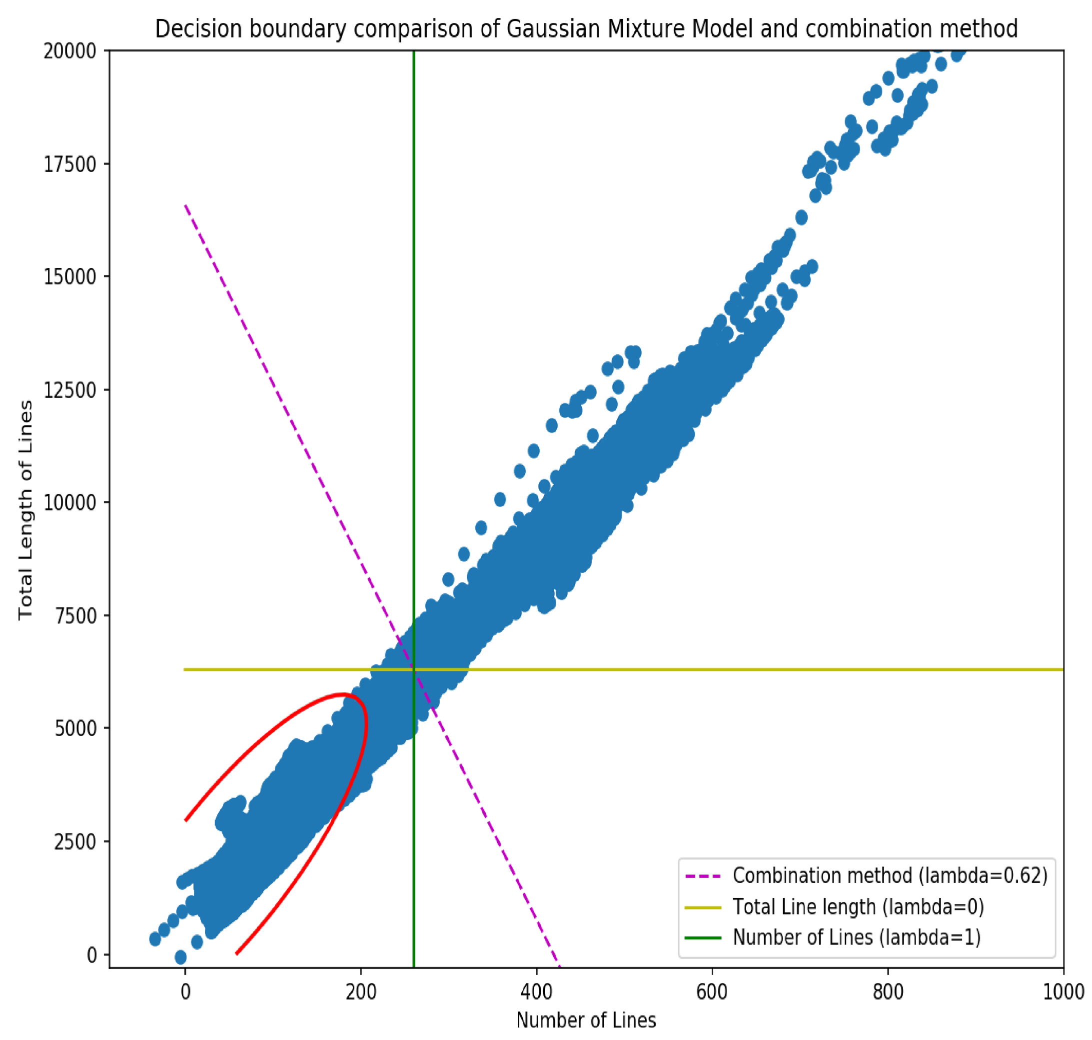
$$Status = \begin{cases} Present \quad \text{if } \lambda \cdot \frac{N.L}{260} + (1-\lambda) \cdot \frac{T.L.L}{6300} \geq 1, \quad where\ N.L : Number\ of\ Lines \\ \text{Absent} \quad \text{if } \lambda \cdot \frac{N.L}{260} + (1-\lambda) \cdot \frac{T.L.L}{6300} < 1 \quad and\ T.L.L : Total\ length\ of\ Lines. \end{cases}$$

Mixture model approach:

- Use GMM on smoothed 2-dimensional feature space to cluster eelgrass from no-eelgrass frames by maximising the likelihood. $\mathcal{L}(\pi,\mu,\Sigma) = \log p(X|\mu,\Sigma,\pi) = \sum_{i=1}^{N} \log p(x_i|\mu,\Sigma,\pi) = \sum_{i=1}^{N} \log \left[\sum_{k=1}^{K} \pi_k \mathcal{N}(x_i|\mu_k,\Sigma_k) \right].$
- Learn the decision boundary of the GMM with equal probability of being eelgrass and no-eelgrass.



Snapshot of final result of the algorithm. A neural network detects the embedded texts in real time.



Comparison of the decision boundary obtained on 2-dimensional feature space from all the methods. GMM has a non-linear boundary and comes close to decision boundary obtained from heuristic method.

Results:

North(start)	East(start)	North(end)	East(end)	Expert	$\lambda = 0$	$\lambda = 0.62$	$\lambda = 1$	GMM
5542.649	1206.088	5542.656	1206.060	0%	0%	0%	0%	0%
5542.654	1206.006	5542.650	1205.985	50%	57.1%	57.1%	52.38%	85.7%
5542.649	1205.979	5542.645	1205.968	90%	100%%	91.6%	75%	100%
5542.615	1205.932	5542.599	1205.916	0%	0%	0%	0%	0%
5541.713	1202.014	5541.729	1201.991	0%	0%	0%	0%	0%
5542.453	1200.592	5542.458	1200.598	95%	100%	100%	97.82%	100%
5542.440	1200.575	5542.445	1200.581	80%	100%	100%	100%	100%
5542.445	1200.581	5542.453	1200.592	85%	100%	100%	93.93%	100%
5542.42	1200.563	5542.428	1200.571	90%	59.45%	48%	55.4%	77%
5542.640	1205.959	5542.635	1205.954	0%	100%	100%	98.14%	100%
5542.393	1200.522	5542.398	1200.528	50%	0%	0%	0%	0%
5542.398	1200.528	5542.404	1200.537	70%	0%	0%	0%	0%

- Counting the ratio of Present frames to Total frames gives the percentage coverage. The GMM over-estimates percentage coverage compared to heuristic method.
- A neural network is used to extract the text information embedded in the frames with an accuracy of 98%.
- It also detects labelling mistakes from domain experts as found out from visual inspection of mismatched predicted frames.

Discussion and future possibilities:

- Multiple cameras can be deployed pointing in different directions to capture more information.
- The *present/absent* status obtained can be extrapolated on a satellite image for large areas using methods like Bayesian classification or Deep learning.

PARTNERS: DHI, AARHUS UNIVERSITET, RAMBØLL, COWI, Miljøstyrelsen Miljø- og Energiministeriet, Vejdirektoratet