Altimetry Waveform Classification and Retracking Strategy for Improved Coastal Altimetry Products

Shubhi Kant Balaji Devaraju

Indian Institute of Technology Kanpur, Uttar Pradesh, India

Sharing is encouraged

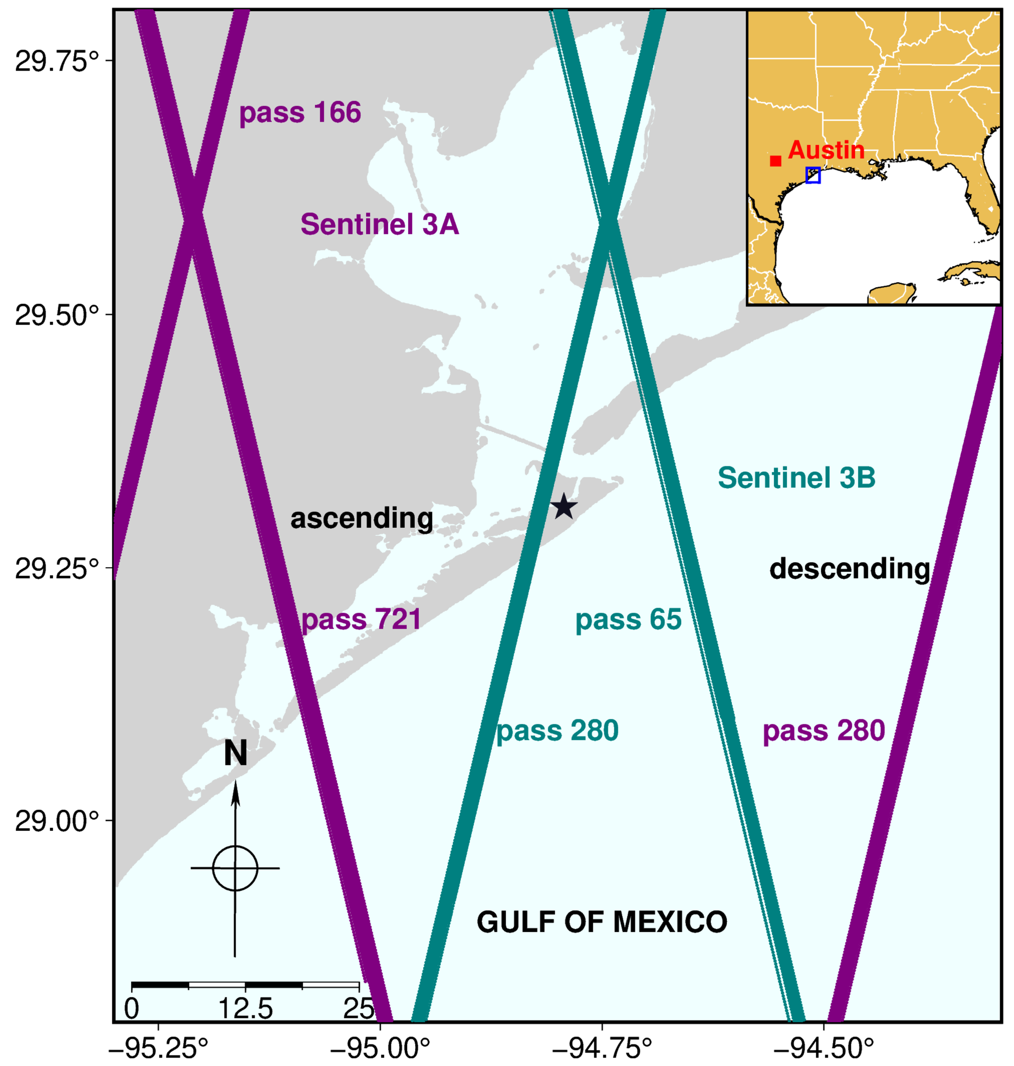
Introduction

Satellite altimetry measures the time taken by a radar pulse to travel to the Earth's surface and back. The two-way travel time measurement is done by estimating the position of the retracking gate from the return power echo. The return waveforms in coastal areas deviate significantly from the typical ocean waveforms, mostly because coastal areas are prone to land interference, varied sea surface behaviours, and diverse topographies.

This study aims to develop an optimum post-processing strategy for coastal waveforms. Our strategy involves classifying the coastal waveforms into various classes based on their shape and identifying the retracking algorithm most suitable for the class.

Study Areas and Datasets

We pursue the classification of Sentinel 3A and 3B, SAR band altimetry data. Our study period is 11th March 2016 to 30th December 2021.



Waveform Classification

- The waveforms are normalized by division with their L2 norms and gaussian noise is added to them.
- LSTM auto-encoder is trained to minimise the Mean Squared Error (MSE) between noisy inputs and outputs.
- K Means algorithm is performed over the encoded features to get initial centroids and clusters.
- Following the approaches presented in Xie et. al [2] and Guo et. al [1], auto-encoder is fine-tuned to cluster the data for 200 epochs.
- During fine tuning the auto-encoder is trained to minimise a combination of entropy loss and Mean Squared Error.

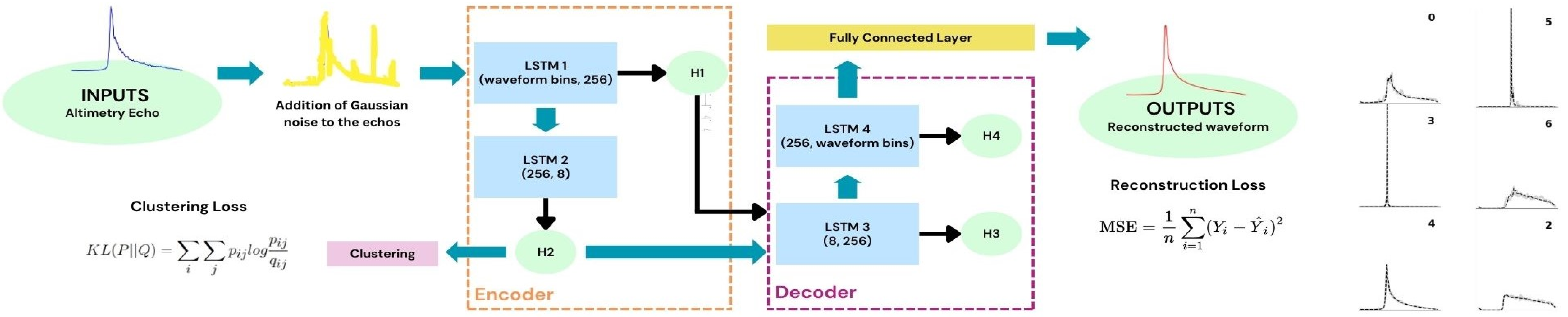
$$KL(P||Q) = \sum_i \sum_j p_{ij} log \frac{p_{ij}}{q_{ij}} \tag{1}$$

where q_{ij} is the similarity between embedded point and cluster center measured using student's t distribution. And p_{ij} is the target distribution defined as

$$p_{ij} = \frac{q_{ij}^2 / \sum_j q_{ij}}{\sum_j (q_{ij}^2 / \sum_j q_{ij})} \tag{2}$$

- During fine tuning the cluster centers are updated at every 5th epoch and it is done till tolerance between old cluster assignments and new cluster assignments reaches 0.0001 or the training stops.

Structure of the Clustering Network



Waveform Classes

Key Observations

- This classification of waveforms captures the spatial distribution of waveforms with class 5 and class 3 located primarily on land areas and class 2 and class4 waveforms located almost exclusively in ocean areas.
- Class 3 which is only there for one cycle, during starting period of Sentinel 3A when it operated in Radar mode instead of SAR mode in coastal areas.
- The waveforms in class 5 and class 3 have multiple maximum peak positions located in different bins. Primarily two distinct clusters are observed with peak lying in early bins (0-50) and peak lying in later bins (> 50)

Discussion

The proposed classification strategy is able to generate distinct waveform classes. This is possible due to learning of invariant representations by auto-encoder. One potential drawback in the present strategy is non-presence of separate classes for multipeak and trash waveforms. This results in potential mixture of these classes with the proposed classes. Since centroid based clustering algorithms tend to generate clusters of equal size, presence of a large number of clusters (25) is observed during clustering. This is due to imbalance in the dataset with class 4 representing more than 50% of waveforms. These clusters are manually grouped back together.

Acknowledgements

Authors would like to thank Copernicus and ESA for making Sentinel 3A and sentinel 3B level 2 datasets used in this study.

References

[1] Xifeng Guo, Long Gao, Xinwang Liu, and Jianping Yin.
Improved deep embedded clustering with local structure preservation.
In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, IJCAI'17, page 1753–1759. AAAI Press, 2017.

[2] Junyuan Xie, Ross Girshick, and Ali Farhadi.
Unsupervised deep embedding for clustering analysis.
In Maria Florina Balcan and Kilian Q. Weinberger, editors, *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 478–487, New York, New York, USA, 20–22 Jun 2016. PMLR.

email: shubhka@iitk.ac.in