Spatio-Temporal Embedding For Algal Bloom Forecasting

TU Delft

Author: Kerem Can Bayraktar K.C.Bayraktar@student.tudelft.nl

Supervisor(s): Jan van Gemert, Attila Lengyel, Robert-Jan Bruintjes

1. BACKGROUND

- "Algal Bloom" is the accumulation of algae in a confined space, and they may harm ecosystems negatively [1]
- Forecasting algal bloom could mitigate the ecological and economical damage they are causing [2]
- Forecasting can be done with deep learning methods through remote sensing data
- Encoded embedding of spatio-temporal information to feature space could improve the performance of deep learning models by capturing spatio-temporal patterns seen in Figure 1

2. RESEARCH QUESTION

Does the inclusion of explicit spatio-temporal embedding methods display a significant improvement for predicting algal blooms?

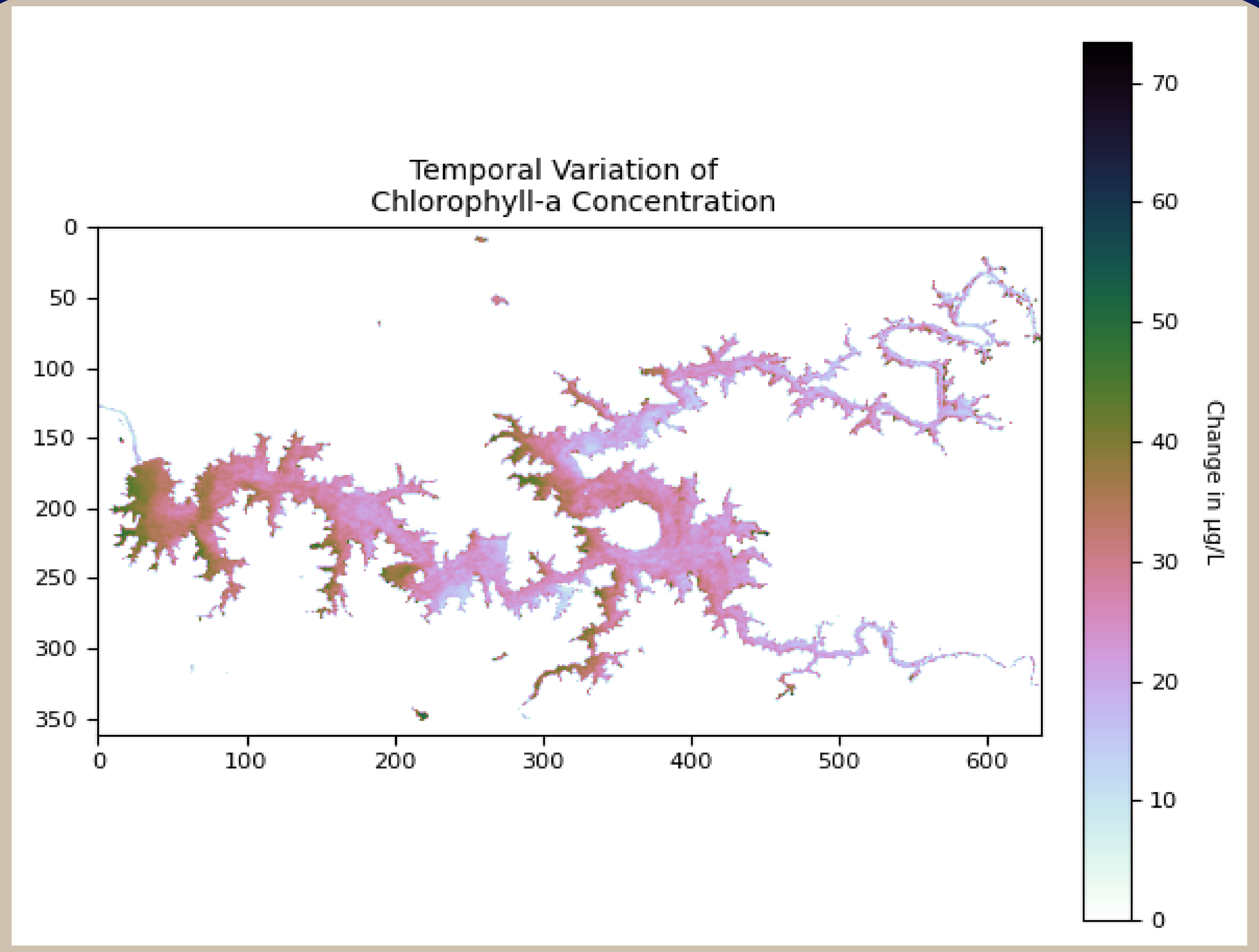

Figure 1. Change in chlorophyll-a values.

3. METHODS

- UNet was the chosen deep learning model to predict chlorophyll-a concentrations by classifying them under bins with intervals [0-10, 10-30, 30-75, 75+] µg/L.
- Input samples of 11 different data types/bands were combined in a data cube seen in Figure 2
- Both the input and ground truth data was clipped and normalized with null values replaced with mean per data type for input and 0 for ground truth
- Ground truth data was labelled into five classes with one background class

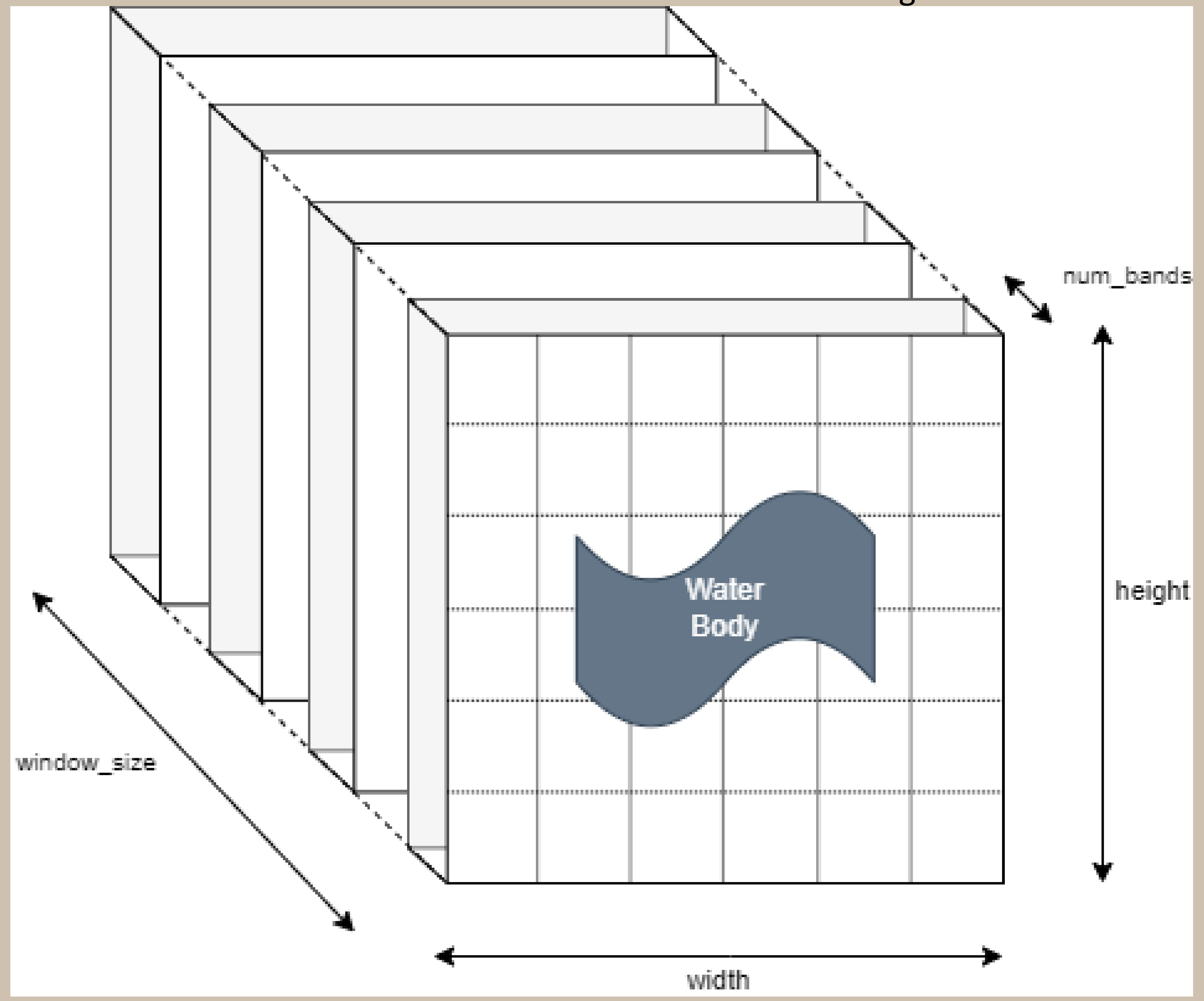

Figure 2. Data cube consisting of combined data modalities from remote sensing data

- Timestamps of individual samples were encoded as day of the year, normalized and added as a new band in order to capture the temporal trends over the dataset's timespan displayed in Figure 1
- Geographic information was extracted from the samples and coordinate values relative to image boundaries were added to each pixel in order to capture the spatial distribution displayed in Figure 1
- The model was trained with spatial, with temporal and with spatio-temporal and without any embedding for comparison

4. RESULTS

embed_type	train_acc	train_loss	val_acc	val_loss
none	0.7668	1.095	0.3336	1.515
spatial	0.7279	0.8591	0.3736	1.467
temporal	0.7384	0.9516	0.375	1.418
spatio-temporal	0.7759	0.8707	0.3305	1.469

Table 1. Respective results with accuracy and weighted cross entropy loss.

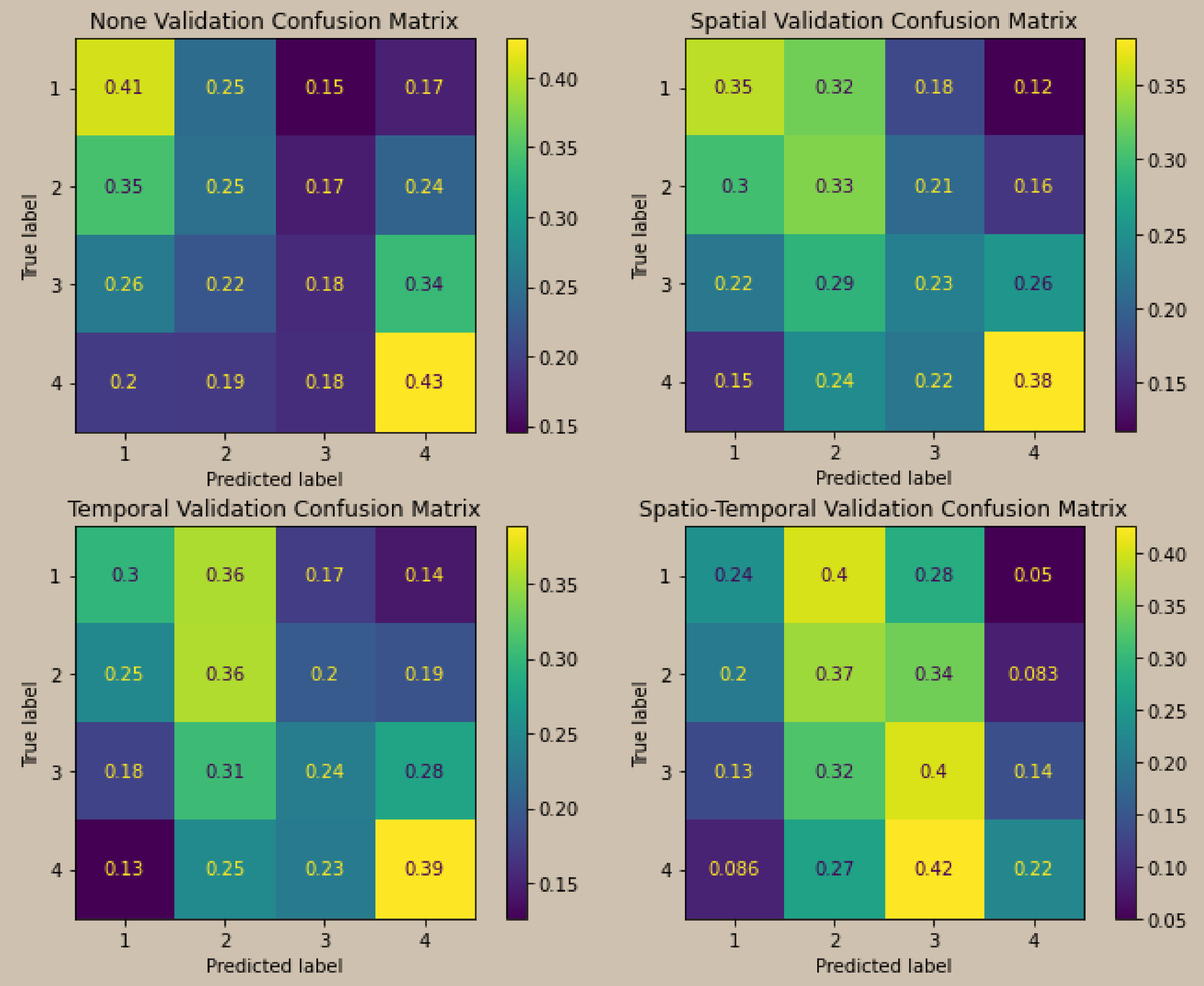

Figure 3. Confusion matrices of respective results, normalized per row.

- The separate inclusion of spatial and temporal embeddings display a decrease in validation losses as well as a slight improvement in validation accuracies(Table 1)
- The combined spatio-temporal embedding shows a decrease in loss although it does not show an improvement in accuracy compared to the one without (Table 1).

5. CONCLUSION

- The results show that the inclusion of explicit spatial and temporal information exhibits a small increase in accuracy. The combined addition of spatio-temporal information did not display a significant improvement for the predictions.
- Figure 3 shows that inclusion of spatial and temporal embeddings allows better predictions for intervals 10-30 and 30-75 than the version without embeddings and less probable to make misclassifications far off the mark (predicting a 0-10 as 75+ and vice versa).
- The accuracy per class observations can be explained by the spatial and temporal patterns forcing the model to predict "conservatively", avoiding costly classifications in exchange of relatively high accuracy percentage for very low and very high intervals.

6. LIMITATIONS & FUTURE WORK

- Conv-LSTM could be used instead to tackle the shortcomings of UNet seen in accuracy and loss scores
- Data imbalance could be dealt with more complex sampling methods and handling of NaN values could be more fitting to match the available data
- Different spatio-temporal encoding methods could be used such as spatial interpolation and wavelet analysis to further explore patterns through the processing of spatio-temporal information

REFERENCES

- [1] A. Mozoet al., "Chlorophyll soft-sensor based on machine learning models for algal bloom predictions,"Scientific Reports, vol. 12, no. 1, 2022.DOI: 10.1038/s41598-022-17299-5.
- [2] J. Heisleret al., "Eutrophication and harmful algal blooms: A scientific consensus,"Harmful Algae,vol. 8, no. 1, pp. 3–13, 2008.DOI: 10.1016/j.hal.2008.08.006.
- [3] X. Shi, Z. Chen, H. Wang, D. Yeung, W. Wong, and W. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," in Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015, December 7-12, 2015, Montreal, Quebec, Canada, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, Eds., 2015, pp. 802–810.