Monitoring Spartina Alterniflora Using Self-Supervised Learning

Abstract

The invasive marsh grass Spartina alterniflora poses a significant threat to coastal ecosystems. This project aimed to develop a robust monitoring system using self-supervised learning techniques applied to satellite imagery data. We explored two methods, SimCLR and DINO, to effectively detect and track the spread of Spartina alterniflora in coastal regions. The key challenges included data sparsity and the lack of annotated datasets, which were addressed through the use of data augmentation and alternative data sources. Our findings demonstrate the potential of self-supervised learning in enhancing the accuracy and generalization of Spartina alterniflora detection, paving the way for improved environmental monitoring and conservation efforts.

Introduction

Invasive species pose a significant threat to the delicate balance of coastal ecosystems worldwide. One such invasive species is the smooth cordgrass, Spartina alterniflora, which has rapidly spread across many coastal regions, outcompeting native vegetation and disrupting the natural habitats (Bortolus, 2006). Effective monitoring and detection of Spartina alterniflora are crucial for managing and mitigating its impact on fragile coastal environments.

Traditionally, the detection and mapping of Spartina alterniflora have relied on manual field surveys or manual interpretation of satellite imagery, which are time-consuming, labor-intensive, and often limited in spatial coverage (Zhong et al., 2012). The emergence of advanced machine learning techniques, particularly self-supervised learning, offers a promising avenue to address these challenges and enhance the accuracy and scalability of Spartina alterniflora monitoring.

The goal of this project was to explore the application of self-supervised learning models, specifically SimCLR (Chen et al., 2020) and DINO (Caron et al., 2021), to detect and monitor the presence and spread of Spartina alterniflora using satellite imagery data. The project aimed to overcome the challenges posed by data sparsity and the lack of annotated datasets, which are common in ecological studies.

Materials and Methodology :

1. Data Collection:
    - We utilized a diverse dataset comprising 2,525 Sentinel-2 RGB images from coastal regions of China and 286 NAIP (National Agriculture Imagery Program) images from coastal Virginia, USA.
    - The data was further enriched by incorporating the GLAD ARD (Global Land Analysis and Discovery Automated Reference Data) dataset, which provided globally consistent Landsat satellite imagery. We selected tiles from the U.S. and China with a high density of Spartina alterniflora.
2. First Method: SimCLR (CNN–Object-based)
    - Pretext Task: We employed stochastic data augmentation techniques, including cropping, flipping, colorization, Gaussian blur, and affine transformation, with a ResNet-50 base encoder.
    - Downstream Task: We used knowledge distillation to incorporate the learned representations into a U-Net model for segmentation of Spartina alterniflora.
    - Observations: The training loss decreased steadily, indicating effective learning, while validation losses showed fluctuations due to data sparsity but an overall improvement in generalization.
3. Second Method: DINO (Self-Supervised Vision Transformer)
    - Pretext Task: We applied self-supervised learning with Vision Transformers (ViT) using stochastic data augmentation.
    - Attention Visualization: We utilized the attention heads from the DINO model to visualize features and detect Spartina alterniflora.
    - Observations: The model exhibited effective learning, as indicated by the decreasing training loss and increasing weight decay, which contributed to improved generalization.

Results and Discussion

1. SimCLR Method:
    - The model showed effective learning with decreasing training loss over epochs. However, the slow pace of learning and fluctuations in validation loss pointed to challenges in using non-annotated Spartina alterniflora data.
2. DINO Method:
    - The attention heads successfully identified Spartina alterniflora in high-resolution NAIP images, and using NIR, Red, and Green bands enhanced vegetation detection. Adding more spectral bands could further improve accuracy.

Conclusion

This project demonstrated the potential of self-supervised learning techniques, such as SimCLR and DINO, in enhancing the detection and monitoring of the invasive marsh grass Spartina alterniflora. Despite the challenges posed by data sparsity and the lack of annotated datasets, the methodologies developed offer promising avenues for further research and improvement.

The experience gained through this project has reinforced our interest in pursuing careers in bioinformatics and environmental conservation, with a focus on the application of machine learning in ecological studies. The insights and knowledge acquired will contribute to the development of more robust and scalable monitoring systems for invasive species, ultimately supporting effective conservation and management strategies.

Acknowledgments

I would like to express our sincere gratitude to our project managers, Keqi He, and project leads, Wenhong Li and Ding Ma, for their guidance and support throughout this project. We also extend our appreciation to our team members, Othmane Echchabi, Advait Bhaskar Pandit, and Meixiang Du, for their valuable contributions and collaborative efforts.

References

- Jackson, M. V., Fuller, R. A., Gan, X., Li, J., Mao, D., Melville, D. S., ... & Choi, C. Y. (2021). Dual threat of tidal flat loss and invasive Spartina alterniflora endanger important shorebird habitat in coastal mainland China. *Journal of environmental management, 278*, 111549.
- O'Donnell, J. P., & Schalles, J. F. (2016). Examination of abiotic drivers and their influence on Spartina alterniflora biomass over a twenty-eight year period using Landsat 5 TM satellite imagery of the Central Georgia Coast. *Remote Sensing*, *8*, 477.
- Roman, C. (2001). Salt Marsh Vegetation. In *Encyclopedia of Ocean Sciences* (pp. 2487–2490). Academic Press. https://doi.org/10.1006/rwos.2001.0088
- Seneca, E. D., Broome, S. W, Woodhouse, W. W., Cammen, L. M., & Lyon, J. T. (1976). Establishing Spartina alterniflora marsh in North Carolina. *Environmental Conservation, 3*, 185-188.
- Zhang, X., Xiao, X., Wang, X., Xu, X., Qiu, S., Pan, L., ... & Li, B. (2023). Continual expansion of Spartina alterniflora in the temperate and subtropical coastal zones of China during 1985-2020. *International Journal of Applied Earth Observation and Geoinformation, 117*, 103192.