Master Thesis Subject:

Learning discrete and interpretable representation of environmental data for wildfire forecasting.

In recent decades, wildfires have increasingly deviated from their natural role in Earth's ecosystems due to climate change and the intensification of fire weather. Terrestrial ecosystems that were once largely unaffected by fires are now under significant threat from large wildfires, such as the boreal ecosystem that was devastated in 2023. Beyond the substantial impact on local ecology, this phenomenon poses a more global threat, as the boreal forest is one of the largest terrestrial carbon sinks [3]. Its destruction exacerbates climate change by releasing stored carbon. Additionally, wildfires represent a major natural hazard to human society, causing extensive material damage and loss of life.

As a result, it is crucial to enhance our understanding of wildfire drivers and improve our ability to accurately forecast wildfire risks on a global scale in the context of climate change. Historically, wildfire forecasting has relied on fire weather indices such as the Canadian Forest Fire Weather Index [18]. However, Machine Learning, particularly Deep Learning, leveraging the growing availability of open-source Earth Observation data [5], holds the potential to surpass traditional process-based approaches that depend on established scientific relationships between fire drivers. Wildfire forecasting can be approached in various ways, including monthly regression of wildfire frequency and size based on tabular environmental data [4, 1], dense predictions of wildfire occurrences at different time horizons incorporating spatial context [11, 15, 20], dense predictions of wildfire spread [12], and next-frame video prediction for wildfire spread [10]. In this project we will focus on the framework from [15] presented in Figure 1 by leveraging the SeasFire dataset [13] containing global wildfire drivers and labels at a spatial resolution of 0.25 degrees.

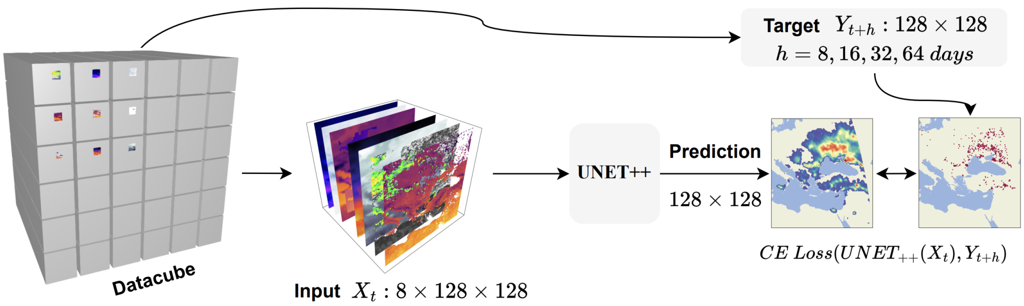

Figure 1: Wildfire forecasting as a dense prediction task [15].

The methods mentioned previously aim to improve wildfire risk prediction but lose the link between environmental drivers and wildfire risk provided by fire weather indices. This project leverages eXplainable Artificial Intelligence (XAI) to enhance wildfire forecasting, aiming to discover and confirm scientific insights on global fire seasons through interpretable model structures. The field of XAI aims to mitigate the risks of model opacity by making aspects of the decision process understandable to humans [8]. XAI methods are divided into post-hoc and by-design approaches: post-hoc methods explain black-box models after training, while by-design methods ensure interpretability within the model

ECEO

EPFL ENAC IIE
Prof. Devis Tuia
Rue de l'industrie 17
Case Postale 440
CH - 1951 Sion

Tel. +41 21 69 382 83 secr.
E-mail. devis.tuia@epfl.ch
E-mail. hugo.porta@epfl.ch

itself.

In this project, we will focus on per-design methods as post-hoc approaches have limitations in terms of faithfulness, detailed information, and completeness [16] despite being commonly used in wildfire forecasting [20]. In particular, we will investigate case-based reasoning methods like part-prototypical networks [6] that discretized the final latent representation space via learned prototypes. This method that targeted image classification was extended to semantic segmentation via parametric [17] and non-parametric prototypes [21]. Figure 2 shows the model architecture for a part-prototypical network [17] in the context of semantic segmentation. The objective of this project is to investigate the interpretable results from existing part-prototypical networks applied to wildfire forecasting and provide a potential extension to those approaches for temporal and geo-reference data. It would be of particular interest to investigate a multi-scale grouping extension of [17] developed at the ECEO lab and iteratively validate the prototypes based on environmental knowledge following [2]. Other classification-based extensions could be investigated such as enforcing orthogonality among prototypes [19], grouping prototypes spatially in a non-rigid way [7], and leveraging learned prototypes in tree-like structure [14]. More generally methods of discrete representation learning [18, 9] that learn similar discrete representation than prototype learning could be tested.

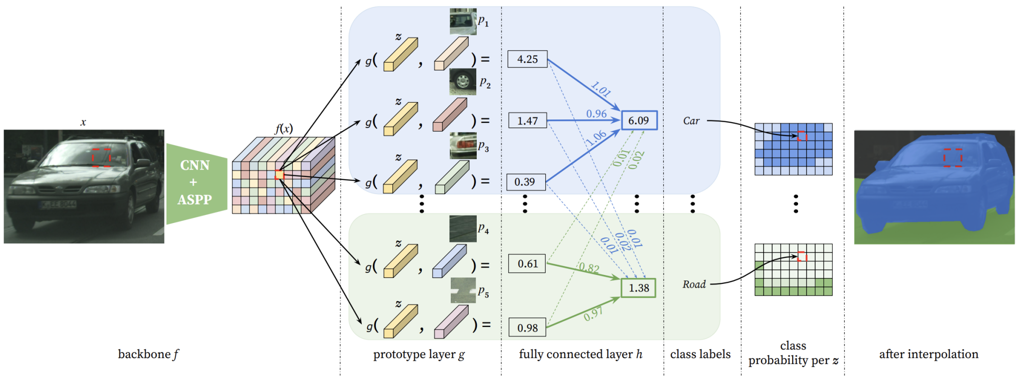

Figure 2: Part-prototypical learning for semantic segmentation. [17].

Requirements:

- Experience in deep learning, especially in computer vision (knowledge of XAI is a plus).
- Proficiency in Python and relevant libraries such as Scikit-learn and Pytorch.
- Familiarity with remote sensing and environmental science is a plus.
- Strong willingness to learn and ability to work independently.

Contact: Prof. Devis Tuia, devis.tuia@epfl.ch

EPFL ENAC IIE
Prof. Devis Tuia
Rue de l'industrie 17
Case Postale 440
CH - 1951 Sion

Tel. +41 21 69 382 83 secr.
E-mail. devis.tuia@epfl.ch
E-mail. hugo.porta@epfl.ch

References:

[1] Sigrid Jørgensen Bakke et al. "A data-driven model for Fennoscandian wildfire danger". In: *Natural hazards and earth system sciences* 23.1 (2023), pp. 65–89.

[2] Andrea Bontempelli et al. "Concept-level debugging of part-prototype networks". In: *arXiv preprint arXiv:2205.15769* (2022).

[3] Corey JA Bradshaw and Ian G Warkentin. "Global estimates of boreal forest carbon stocks and flux". In: *Global and Planetary Change* 128 (2015), pp. 24–30.

[4] Jatan Buch et al. "SMLFire1. 0: a stochastic machine learning (SML) model for wildfire activity in the western United States". In: *Geoscientific Model Development* 16.12 (2023), pp. 3407–3433.

[5] Gustau Camps-Valls et al. *Deep learning for the Earth Sciences: A comprehensive approach to remote sensing, climate science and geosciences*. John Wiley & Sons, 2021.

[6] Chaofan Chen et al. "This looks like that: deep learning for interpretable image recognition". In: *Advances in neural information processing systems* 32 (2019).

[7] Jon Donnelly, Alina Jade Barnett, and Chaofan Chen. "Deformable protopnet: An interpretable image classifier using deformable prototypes". In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022, pp. 10265–10275.

[8] Finale Doshi-Velez and Been Kim. "Towards a rigorous science of interpretable machine learning". In: *arXiv preprint arXiv:1702.08608* (2017).

[9] Srishti Gautam et al. "Protovae: A trustworthy self-explainable prototypical variational model". In: *Advances in Neural Information Processing Systems* 35 (2022), pp. 17940–17952.

[10] Thai-Nam Hoang, Sang Truong, and Chris Schmidt. "Wildfire Forecasting with Satellite Images and Deep Generative Model". In: *arXiv preprint arXiv:2208.09411* (2022).

[11] Fantine Huot et al. "Deep learning models for predicting wildfires from historical remote-sensing data". In: *arXiv preprint arXiv:2010.07445* (2020).

[12] Fantine Huot et al. "Next day wildfire spread: A machine learning dataset to predict wildfire spreading from remote-sensing data". In: *IEEE Transactions on Geoscience and Remote Sensing* 60 (2022), pp. 1–13.

[13] Ilektra Karasante et al. "SeasFire as a Multivariate Earth System Datacube for Wildfire Dynamics". In: *arXiv preprint arXiv:2312.07199* (2023).

[14] Meike Nauta, Ron Van Bree, and Christin Seifert. "Neural prototype trees for interpretable fine-grained image recognition". In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2021, pp. 14933–14943.

[15] Ioannis Prapas et al. "Deep learning for global wildfire forecasting". In: *arXiv preprint arXiv:2211.00534* (2022).

[16] Cynthia Rudin. "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead". In: *Nature machine intelligence* 1.5 (2019), pp. 206–215.

[17] Mikołaj Sacha et al. "Protoseg: Interpretable semantic segmentation with prototypical parts". In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2023, pp. 1481–1492.

ECEO

EPFL ENAC IIE
Prof. Devis Tuia
Rue de l'industrie 17
Case Postale 440
CH - 1951 Sion

Tel. +41 21 69 382 83 secr.
E-mail. devis.tuia@epfl.ch
E-mail. hugo.porta@epfl.ch

[18] Charles E Van Wagner et al. *Structure of the Canadian forest fire weather index*. Vol. 1333. Environment Canada, Forestry Service Ottawa, ON, Canada, 1974.

[19] Jiaqi Wang et al. “Interpretable image recognition by constructing transparent embedding space”. In: *Proceedings of the IEEE/CVF international conference on computer vision*. 2021, pp. 895–904.

[20] Guoli Zhang, Ming Wang, and Kai Liu. “Deep neural networks for global wildfire susceptibility modelling”. In: *Ecological Indicators* 127 (2021), p. 107735.

[21] Tianfei Zhou et al. “Rethinking semantic segmentation: A prototype view”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022, pp. 2582–2593.

ECEO

EPFL ENAC IIE
Prof. Devis Tuia
Rue de l’industrie 17
Case Postale 440
CH - 1951 Sion

Tel. +41 21 69 382 83 secr.
E-mail. devis.tuia@epfl.ch
E-mail. hugo.porta@epfl.ch