Expert-in-the-loop Systems Towards Safety-critical Machine Learning Technology in Wildfire Intelligence

Maria João Sousa[*], **Alexandra Moutinho**
IDMEC, Instituto Superior Técnico,
Universidade de Lisboa
maria.joao.sousa@tecnico.ulisboa.pt
alexandra.moutinho@tecnico.ulisboa.pt

Miguel Almeida
Forest Fire Research Center,
ADAI, University of Coimbra
miguelalmeida@adai.pt

Abstract

With the advent of climate change, wildfires are becoming more frequent and severe across several regions worldwide. To prevent and mitigate its effects, wildfire intelligence plays a pivotal role, *e.g.* to monitor the evolution of wildfires and for early detection in high-risk areas such as wildland-urban-interface regions. Recent works have proposed deep learning solutions for fire detection tasks, however the current limited databases prevent reliable real-world deployments. We propose the development of expert-in-the-loop systems that combine the benefits of semi-automated data annotation with relevant domain knowledge expertise. Through this approach we aim to improve the data curation process and contribute to the generation of large-scale image databases for relevant wildfire tasks and empower the application of machine learning techniques in wildfire intelligence in real scenarios.

1 Introduction

Wildfires are a recurrent natural hazard on a global scale that has a brutal impact on the environment and natural ecosystems, which can lead to disasters with dire impacts on communities [1]. As a result of climate change, fire events are becoming more frequent and severe, with meteorological conditions of high ignition propensity being more frequently met, leading to increased fire spotting, and rapid spread. In addition, these conditions are verified over longer periods, extending fire seasons in several regions worldwide.

To prevent and mitigate the devastating effects of wildfire events, it becomes urgent to detect fires in an early stage and to monitor wildfires in near-real-time as the events unravel, providing enhanced situational awareness for decision-making and operational teams. In that sense, wildfire intelligence plays a pivotal role, especially for high-risk areas such as wildland-urban-interface regions [2] and large-scale wildfires [3]. For these reasons, there is a current demand for improvements and increased levels of automation in the stages of pre-fire event, firefighting, and aftermath.

In this context, the breakthroughs in machine learning (ML) can be an enabling technology towards the integration of artificial intelligence products in current decision support systems. More specifically, related works have proposed ML solutions for image-based fire detection tasks [4, 5], however the quality and limited size of image databases available often do not offer generalization guarantees for reliable deployments in real contexts [6]. Although transfer learning and data augmentation techniques have been explored in related work, the limitations in interpretability and transparency of black-box models prevent effectively fine-tuning these models to solve existing shortcomings [6]. The lack of large-scale databases for wildfire detection and monitoring tasks is a known hurdle in developing machine learning algorithms with adequate generalization.

[*]Corresponding author.

Tackling Climate Change with Machine Learning workshop at NeurIPS 2020.

Considering that employing ML solutions for wildfire intelligence involves deployments in safety-critical applications, the robustness and reliability of the models have yet to undergo significant developments. Conversely, this may be achieved through high-quality data curation, despite not alleviating black-box limitations, or through the exploration of algorithms with increased explainability, fine-tuning ability and interpretability. However, both these scenarios call for the input of expert knowledge to develop accurately annotated data, which is particularly nuanced in the field of wildfire management and operations, and requires domain expertise as there is also high data uncertainty.

To address these issues, we propose the development of expert-in-the-loop systems that combine the automation advantages provided by machine learning with the introduction of relevant domain knowledge expertise. The main objective of the proposed approach is the development of semi-automated software tools that can support data curation by wildfire experts. This solution can allow a better handling of data uncertainty and improve the quality of data sources, with the ultimate objective of enabling the development of large-scale databases for wildfire-related problems. More importantly the expert-in-the-loop approach allows involving domain experts and end-users (*e.g.,* researchers and public agencies) in the development procedure, thus improving the relevance of machine learning applications developed for wildfire intelligence in real contexts.

2 Image-based Wildfire Management Tasks

Image data for wildfire-related tasks can have a broad spectrum of characteristics and modalities as these can be collected *e.g.*, from satellites, aerial vehicles, watchtowers, or ground teams. The latency associated with these data types also varies and consequently, so does its timescales of application. From this plurality arises a great breadth of opportunities for ML approaches to deal with high-dimensional data, but also a great challenge for data curation. Herein, we outline several relevant tasks followed by a brief data description.

Tasks Wildfire management involves four main stages: *i)* prevention, *ii)* preparedness, *iii)* response and *iv)* recovery. The following tasks can extend to several of these stages and involve processing of large amounts of image data, thus having a high potential interest and benefit from the usage of ML.

- Risk assessment concerning environmental conditions and risk mapping based on land use and social patterns;
- Vegetation management to reduce fire severity, *e.g.* fuel mapping, monitoring of fuel breaks, or tracking of vegetation fuel moisture content;
- Wildfire detection and monitoring, *e.g.*, early identification of flames and smoke plume, mapping of the firefront(s), early detection of spot fires and identification of hot spots;
- Post-event analyses, *e.g.*, mapping burned areas, evaluation of possible subsequent cascading effects, e.g., erosion risks, and air quality estimation based on remote sensing.

Data The evolving datasets being developed comprise multimodal image data, currently in the visible and thermal infrared bands, including images captured from ground teams, watchtowers, aircraft and high-altitude balloons, which are exemplified in Fig. 1 (for visible range instances only). The image samples comprise a diverse collection of situations, with most concerning real wildfire events and field experimental burns. To create a balanced and robust set of data, factors that may induce misclassifications are also included such as clouds, fog or sunsets, as well as firefighting vehicles, power lines and various types of operational teams.

Figure 1: Samples of fire and not fire instances captured from ground teams, drones and high-altitude balloons, along with aerial vehicles and teams in operational missions in a wide variety of scenarios and lighting conditions.

3 Proposed Approach

The development of expert-in-the-loop systems aims to bridge the gap between machine learning automation (*e.g.* in classification, segmentation, or detection tasks) and the inclusion of relevant domain knowledge expertise, so that data curation is relevant for wildfire intelligence in real-world scenarios and wildfire science research. Previous contributions in the literature have favored techniques that are well-suited for real-time deployments despite having limitations in performance, in lack of transparency and interpretability. However, those limitations hamper considerably the reliability and acceptance of end-users of such techniques, hindering the deployment in safety-critical applications in real contexts.

This novel approach in wildfire related applications leverages the potential of exploring machine learning and computer vision, and intelligent systems methods that have not been particularly designed for online computational performance (in terms of speed/energy) in real-time deployments, but are rather accurate despite computationally heavy and/or time-consuming. To harness the advantages of this approach, this project aims to design expert-in-the-loop systems and develop software tools that can introduce domain knowledge into the data curation and task design processes. To that end, this approach can be outlined as exemplified in the diagram in Fig. 2. Domain expertise is introduced in two stages through two feedback loops: 1) task definition/refinement - where experts define relevant tasks and refine these based on the results of the data curation process (outer loop); and 2) expert validation and interpretation - where at the end of automated processing pipeline, experts enable re-iteration and learning patterns on verified outputs (inner loop). Depending on the task at hand, feature extraction, feature selection, detection and pixel-level segmentation techniques used for semi-automated annotation can resort to several computer vision, machine learning and a broad scope of intelligent systems approaches. Successful implementations of this approach should aim for obtaining increasingly accurate fine-grained outputs validated with cohorts of experts, which shall be quantified with relevant evaluation metrics for quantitative benchmarking.

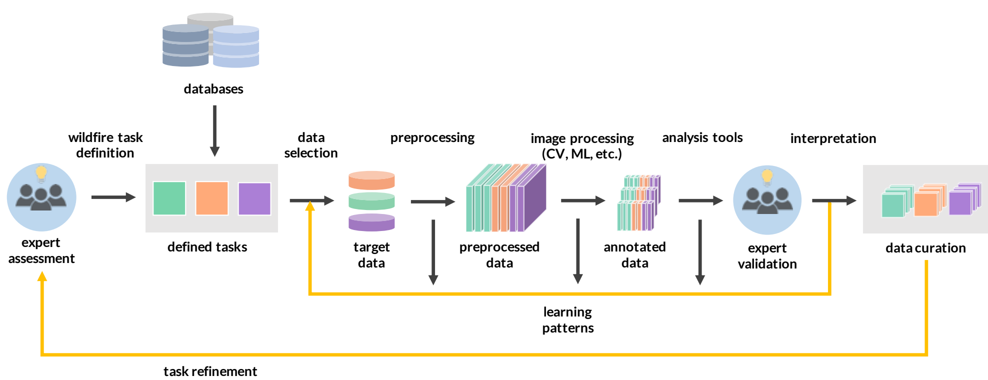

Figure 2: Expert-in-the-loop system comprising a computational data annotation pipeline with expert feedback.

The key benefit of this data curation approach is to yield fine-grained annotated data sources for the development of large-scale databases for wildfire-related problems. By being validated by domain experts, it will also improve the relevance of subsequent machine learning applications developed for end-users (*e.g.* wildfire management, firefighting and civil protection agencies, and researchers working on wildfire-related topics).

4 Conclusion and Future Work

The solutions developed through the proposed approach will be an important stepping stone for data curation and creating large-scale datasets for wildfire detection and monitoring. These datasets will open opportunities for leveraging ML technologies in this context, as well as pave the way for relevant multi-view data association [7] and autonomous robotics tasks in this domain [8]. ML along with emerging technologies such as unmanned aerial vehicles and cube-sat systems are important enabling technologies for near-real-time wildfire intelligence, which will have an essential role in decision support systems with crucial impacts in the safety of at risk populations and environment protection.

Broader Impact

Real-time early fire detection and monitoring systems can prevent the loss of natural ecosystems responsible for climate regulation through carbon sequestration, can help to avoid the occurrence of large burnt areas, and the emission of greenhouse gases. Therefore, the integration, at different timescales, of data-driven intelligent systems in decision support systems for firefighting and civil protection can contribute to mitigate the social, cultural, environmental and economic effects associated with wildfires, contributing to the United Nations Sustainable Development Goals.

Acknowledgments and Disclosure of Funding

This work was financed by national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., through IDMEC under project Eye in the Sky (PCIF/SSI/0103/2018), and through IDMEC under LAETA, project UIDB/50022/2020. M.J. Sousa acknowledges the support from FCT through the Ph.D. scholarship SFRH/BD/145559/2019.

References

[1] Douglas Paton, Petra T. Buergelt, Sarah McCaffrey, and Fantina Tedim, editors. *Wildfire Hazards, Risks, and Disasters*. October 2014.

[2] David E Calkin, Jack D Cohen, Mark A Finney, and Matthew P Thompson. How risk management can prevent future wildfire disasters in the wildland-urban interface. *Proceedings of the National Academy of Sciences*, 111(2):746–751, 2014.

[3] Fantina Tedim, Vittorio Leone, Malik Amraoui, Christophe Bouillon, Michael R Coughlan, Giuseppe M Delogu, Paulo M Fernandes, Carmen Ferreira, Sarah McCaffrey, Tara K McGee, et al. Defining extreme wildfire events: difficulties, challenges, and impacts. *Fire*, 1(1):9, 2018.

[4] Piyush Jain, Sean C.P. Coogan, Sriram Ganapathi Subramanian, Mark Crowley, Stephen W. Taylor, and Mike D. Flannigan. A review of machine learning applications in wildfire science and management. *Environmental Reviews*, 2020. doi: 10.1139/er-2020-0019.

[5] Anshul Gaur, Abhishek Singh, Anuj Kumar, Ashok Kumar, and Kamal Kapoor. Video flame and smoke based fire detection algorithms: A literature review. *Fire Technology*, 2020.

[6] Maria João Sousa, Alexandra Moutinho, and Miguel Almeida. Wildfire detection using transfer learning on augmented datasets. *Expert Systems with Applications*, 142:112975, 2020.

[7] Deepti Pachauri, Risi Kondor, and Vikas Singh. Solving the multi-way matching problem by permutation synchronization. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 26*, pages 1860–1868. Curran Associates, Inc., 2013.

[8] J. R. Martínez-de Dios, Luis Merino, Aníbal Ollero, Luis M. Ribeiro, and Xavier Viegas. *Multi-UAV Experiments: Application to Forest Fires*, pages 207–228. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007. ISBN 978-3-540-73958-6. doi: 10.1007/978-3-540-73958-6_8.