EO4WildFires: An Earth Observation multi-sensor, time-series machine-learning-ready benchmark dataset for wildfire impact prediction

Dr. Dimitrios Sykas*[a], Dimitrios Zografakis[a], Konstantinos Demestichas[a], Constantina Costopoulou[a], Pavlos Kosmidis[b]
[a] Department of Agricultural Economy & Development of the Agricultural University of Athens – dsykas@aua.gr, [b]Catalink Ltd

ABSTRACT

This paper presents a benchmark dataset called EO4WildFires; a multi-sensor (multi spectral; Sentinel-2, Synthetic-Aperture Radar - SAR; Sentinel-1, meteorological parameters; NASA Power) time-series dataset that spans 45 countries, which can be used for developing machine learning and deep learning methods targeted for the estimation of the area that a forest wildfire might cover.

This novel EO4WildFires dataset is annotated using EFFIS (European Forest Fire Information System) as forest fire detection and size estimation data source. A total of 31,730 wildfire events are gathered from 2018 to 2022. For each event, Sentinel-2 (multispectral), Sentinel-1 (SAR) and meteorological data are assembled into a single data cube. The meteorological parameters that are included in the data cube are: ratio of actual partial pressure of water vapor to the partial pressure at saturation, average temperature, bias corrected average total precipitation, average wind speed, fraction of land covered by snowfall, percent of root zone soil wetness, snow depth, snow precipitation, as well as percent of soil moisture.

The main problem that this dataset is designed to address, is the severity forecasting before wildfires occur. The dataset is not used to predict wildfire events, but rather to predict the severity (size of area damaged by fire) of a wildfire event, if that happens in a specific place under the current and historical forest status, as recorded from multispectral and SAR images, and meteorological data.

Using the data cube for the collected wildfire events, the EO4WildFires dataset is used to realize three (3) different preliminary experiments, to evaluate the contributing factors for wildfire severity prediction. The first experiment evaluates wildfire size using only the meteorological parameters, the second one utilizes both the multispectral and SAR parts of the dataset, while the third exploits all dataset parts. In each experiment, machine learning models are developed, and their accuracy is evaluated. The results show that the size of wildfire events can be estimated better using Sentinel-2 data. Second in terms of accuracy is Sentinel-1, while the usage of only meteorological data presented the lowest accuracy among the three.

The dataset is published with an Open Access license and is hosted at: 10.5281/zenodo.7762564.

Keywords: forest wildfires, earth observation, machine learning dataset, multi-sensor data

1. INTRODUCTION

Forest wildfires constitute a pressing global concern, causing widespread environmental, economic, and social damage. The increasing frequency and intensity of these events have been attributed to factors such as climate change, deforestation, and human activities. Accurate and timely prediction of wildfire severity, particularly the size of the area that a fire might cover, is crucial for effective disaster management, resource allocation, and mitigation efforts. Remote sensing data, especially from satellite platforms such as Sentinel-1 and Sentinel-2, as well as meteorological data from sources like NASA Power or European Centre for Medium-Range Weather Forecasts (ECMWF), provide a wealth of information that can be harnessed to develop machine learning and deep learning models for wildfire severity estimation.
In this paper, we introduce a novel benchmark dataset called EO4WildFires, which combines multispectral (Sentinel-2), Synthetic-Aperture Radar - SAR (Sentinel-1), and meteorological data from 30 countries to create a comprehensive, multi-sensor time-series dataset. Annotated using the European Forest Fire Information System (EFFIS) for wildfire detection

Ninth International Conf. on Remote Sensing and Geoinformation of the Environment (RSCy2023), edited by K. Themistocleous, et al., Proc. of SPIE Vol. 12786, 1278603 · · doi: 10.1117/12.2680777

and size estimation, EO4WildFires spans 31,742 wildfire events between 2018 and 2022. By assembling Sentinel-2, Sentinel-1, and meteorological data into a single data cube for each event, this dataset allows for a more in-depth analysis of the factors contributing to wildfire severity.
This paper aims to address the challenge of forecasting forest wildfire severity before the events occur, focusing on the potential size of the area damaged by fire in a specific location given the current and historical forest status. To this end, we design three experiments to evaluate the contribution of various factors in predicting wildfire size using the EO4WildFires dataset. Each experiment explores different combinations of meteorological parameters, multispectral, and SAR data, and assesses the accuracy of machine learning models developed based on these inputs. Through these experiments, we aim to provide insights into the development of more accurate and reliable wildfire severity prediction models that can better inform decision-makers and support wildfire management efforts worldwide.

1.1 Deep Learning Architectures

Deep learning technologies are evolving at fast pace. This sub-section provides a brief overview of recent notable deep learning architectures that make use of large-scale visual data.

The authors in [1] introduce the technique of Contrastive Captioner (CoCa), a novel image-text encoder-decoder foundation model that combines contrastive loss and captioning loss. This approach incorporates elements of both CLIP and Simple Visual Language Model (SimVLM). CoCa is designed to first focus on unimodal text representations and then on multimodal image-text representations. The model uses contrastive loss between unimodal embeddings and captioning loss on multimodal decoder outputs. CoCa is pretrained end-to-end on web-scale alt-text data and annotated images, unifying natural language supervision for representation learning. It achieves state-of-the-art performance on various tasks, including visual recognition, cross-modal retrieval, multimodal understanding, and image captioning. Notably, CoCa reaches 91.0% top-1 accuracy on ImageNet with a finetuned encoder.

The authors in [2] explore large-scale models in computer vision and address three major challenges: training instability, resolution gaps, and dependence on labeled data. The authors propose three techniques: 1) a residual-post-norm method with cosine attention for training stability, 2) a log-spaced continuous position bias method to transfer models between different resolution tasks, and 3) SimMIM, a self-supervised pretraining method that reduces the need for labeled images. The paper successfully trains a 3 billion-parameter Swin Transformer V2 model, the largest dense vision model to date, capable of handling images up to 1,536x1,536 resolution. This model sets new performance records on ImageNet-V2, COCO, ADE20K, and Kinetics-400 tasks. Moreover, the training is 40 times more efficient in terms of labeled data and training time compared to Google's billion-level visual models.

The study in [3] demonstrates that the Transformer architecture, which has become the standard for natural language processing tasks, can also be effectively applied to computer vision without relying on convolutional networks. The authors introduce Vision Transformer (ViT), a pure transformer applied directly to sequences of image patches. When pre-trained on large datasets and transferred to various image recognition benchmarks (ImageNet, CIFAR-100, VTAB, etc.), ViT achieves excellent results compared to state-of-the-art convolutional networks while requiring significantly fewer computational resources to train.

A method for algorithm discovery as program search, focusing on optimization algorithms for deep neural network training is presented in [4]. The authors introduce a technique called Evolved Sign Momentum and codenamed Lion, a simple, memory-efficient optimization algorithm that outperforms widely used optimizers like Adam and Adafactor in various tasks. On image classification, Lion improves ViT's accuracy by up to 2% on ImageNet and reduces pre-training compute on JFT datasets by up to 5x. In vision-language contrastive learning, Lion achieves 88.3% zero-shot and 91.1% fine-tuning accuracy on ImageNet. For diffusion models, it outperforms Adam by achieving a better Fréchet inception distance (FID) score and reducing training compute by up to 2.3x. Lion performs similarly or better compared to Adam in autoregressive, masked language modeling, and fine-tuning tasks. Its performance gain grows with larger training batch sizes and requires a smaller learning rate due to the larger norm of the update produced by the sign function.

1.2 Wildfires and Earth Observation

Satellites have a role to play in detecting, monitoring and characterizing fires [5]. This sub-section provides an overview of notable recent research studies that produce or utilize Earth Observation data for purposes of wildfire detection.

The authors in [6] compare MODIS fire products [5] with ground wildfire investigation records from December 2002 to November 2015 in Yunnan Province, Southwest China. The goal is to understand the differences in spatiotemporal patterns of regional wildfires detected by the two approaches, estimate the omission error of MODIS fire products, and explore how local environmental factors influence MODIS wildfire detection probability. The results show that MODIS records at least twice as many wildfire events compared to ground records, but the distribution patterns are inconsistent. Only 11.10% of the 5,145 confirmed ground records were detected using multiple MODIS fire products. The study found that fire size is a primary limiting factor for MODIS fire detection capacity, with a 50% probability of detecting wildfires at least 18 hectares in size. Other factors influencing MODIS wildfire detection probability include weather factors (daily relative humidity and wind speed) and altitude of wildfire occurrence. The study highlights the importance of considering local conditions and ground inspection in wildfire monitoring, management, and global wildfire simulations.

The authors in [7] present a method that combines Big Data, Remote Sensing, and Data Mining algorithms (Artificial Neural Network and Support Vector Machines) to process data from satellite images and predict wildfire occurrences. The authors create a dataset based on Remote Sensing data related to crop conditions (Normalized Difference Vegetation Index - NDVI), meteorological conditions (Land Surface Temperature – LST), and the fire indicator “Thermal Anomalies” from the MODIS instrument on Terra and Aqua satellites. The dataset is publicly available on GitHub. Experiments were conducted using the big data platform “Databricks,” achieving high prediction accuracy (98.32%). The results were assessed using various validation strategies and compared with existing wildfire early warning systems.

The authors in [8] present "Next Day Wildfire Spread", a large-scale, multivariate dataset of historical wildfires in the United States, based on nearly a decade of remote-sensing data. Unlike existing fire datasets, this dataset combines 2-D fire data with multiple explanatory variables (topography, vegetation, weather, drought index, and population density) over 2-D regions, creating a feature-rich dataset for machine learning. The authors implement a neural network that leverages the spatial information in the data to predict wildfire spread, comparing its performance with logistic regression and random forest models. This dataset serves as a benchmark for developing remote-sensing-based wildfire propagation models with a lead time of one day.

The authors in [9] propose a cost-effective, machine-learning-based approach to predict forest fires in Indonesia using remote sensing data. This addresses the challenges faced by developing countries that cannot afford expensive ground instruments used in traditional prediction systems. The proposed model achieves over 0.81 area under the receiver operator characteristic (ROC) curve, significantly outperforming the baseline approach, which never exceeds 0.70 area under the ROC curve. The model maintains its performance even with reduced data, demonstrating the potential for machine-learning-based approaches to create reliable and cost-effective forest fire prediction systems.

2. METHODOLOGY

2.1 Scientific Problem

The scientific problem addressed by the EO4WildFires dataset is to enable the development of a model or set of models for forecasting the severity of a future wildfire event in a specific location, based on current and past (30 days) meteorological data, multispectral and SAR images. These data should model the forest status before a wildfire event occurs. The goal is not to predict wildfire events, but rather to predict the severity, specifically the size of the area damaged by the fire, before it occurs. This problem requires the development of predictive models that can integrate the various types of data available in the dataset to generate accurate severity forecasts. The solution to this problem has the potential to help forest protection services and other stakeholders to better prepare for and respond to wildfires, ultimately reducing their impact on the environment and society.

As a scientific question, it can be articulated as follows:

“What is the potential of the provided dataset, which includes meteorological data, multispectral and SAR images, to forecast the severity (size of area damaged by fire) of a future wildfire event in a specific location, based on current and near past forest status?”

To address this, the European Forest Fire Information System (EFFIS), Copernicus Sentinel-1 & 2, and NASA Power data sources are compiled in a data cube format to enable deep learning modeling.

2.2 European Forest Fire Information System

EFFIS is a platform that provides up-to-date information on wildland fires in Europe to support forest protection services. The website is maintained by the European Commission Joint Research Centre (JRC) and provides data on historical and current wildfires, as well as forecasts of wildfire danger levels. The information on EFFIS includes maps, data on the location, size, and intensity of wildfires, as well as the affected vegetation type and land use. EFFIS also provides a daily fire danger forecast based on meteorological data and a fire danger rating system. The website is an essential resource for forest protection services, researchers, and policymakers, who rely on accurate and timely information to monitor and manage wildfires in Europe.

2.3 Copernicus Sentinel 1 & 2

Sentinel-1 is an Earth observation mission from the Copernicus Programme, developed by the European Space Agency (ESA). It consists of a constellation of two polar-orbiting satellites that systematically acquire Synthetic Aperture Radar (SAR) imagery at high spatial resolution over land and coastal waters. The SAR sensor operates in the C-band frequency and provides all-weather and day-and-night imaging capabilities, making it useful for a wide range of applications, including monitoring of land and ocean surfaces, disaster management, and maritime surveillance. The data from Sentinel-1 is freely available and can be accessed through the Sentinel Data Hub or through third-party data providers. The mission has been designed to provide long-term, global data acquisition, making it an essential tool for monitoring and managing natural resources and the environment.

Sentinel-2 is an Earth observation mission from the Copernicus Programme, developed by the European Space Agency (ESA). It consists of a constellation of two polar-orbiting satellites that systematically acquire optical imagery at high spatial resolution over land and coastal waters. The sensor on board the satellites captures images in 13 spectral bands, ranging from the visible to the shortwave infrared regions of the electromagnetic spectrum. This enables the detection and monitoring of a wide range of phenomena, including vegetation cover, land use, natural disasters, and urban development. The data from Sentinel-2 is freely available and can be accessed through the Sentinel Data Hub or through third-party data providers. Sentinel-2 has been designed to provide global coverage and high revisit time, making it a valuable tool for environmental monitoring, land use mapping, and disaster management.

2.4 NASA Power

NASA Power is a scientific resource that provides solar and meteorological data sets from NASA research to support various applications, such as renewable energy, building energy efficiency, and agricultural needs. The scientific problem addressed by NASA Power is to make solar and meteorological data more accessible and usable for researchers, policymakers, and practitioners working in different fields. This problem requires the development of reliable and accurate methods for measuring and predicting solar and meteorological variables, such as solar radiation, temperature, precipitation, and wind speed. NASA Power provides users with access to a range of data sets, tools, and services that can be used to support research, planning, and decision-making related to energy, agriculture, and other sectors. The solutions provided by NASA Power have the potential to contribute to a more sustainable and resilient future by enabling informed decisions based on accurate and up-to-date solar and meteorological data.

2.5 Dataset Structure

The dataset is a combination of meteorological data, multispectral and SAR satellite images, and wildfire event labels provided by the EFFIS system. The EFFIS system is responsible for providing updated and reliable information on wildland fires in Europe to support forest protection services.

The dataset was created using the Sentinel-hub API and the NASA Power API. The Sentinel-hub API allows users to make Web Map Service (WMS) and Web Coverage Service (WCS) web requests to download and process satellite images from various data sources. The NASA Power API provides solar and meteorological data sets from NASA research to support renewable energy, building energy efficiency, and agricultural needs.

Table 1. Parameters included in each wildfire event for each data source

Meteorological Data	**Sentinel-1**	**Sentinel-2**

ratio of actual partial pressure of water vapor to the partial pressure at saturation (RH2M)	VV	Band 02
average temperature (T2M)	VH	Band 03
bias corrected average total precipitation (PRECTOTCORR)	(VV-VH)/(VV+VH)	Band 04
average wind speed (WS2M)		Band 05
fraction of land covered by snowfall (PRECSNOLAND)		Band 08
percent of root zone soil wetness (GWETROOT)		Band 11
snow depth (SNODP)		
snow precipitation (FRSNO)		
soil moisture (GWETTOP)		

In our produced dataset, EO4WildFires, each wildfire event is packaged as a datacube in NetCDF format, resulting in a total of 31,740 files, equal to the number of wildfire events in the dataset. The data for each wildfire event are collected as follows:

- Bounding box coordinates and the date of the event are used as inputs.
- Meteorological parameters are extracted using the center of the area as the coordinate. Parameters are collected from one day before the event until 30 days earlier than that date.
- Sentinel-2 images are subset using the bounding box coordinates. A mosaicking process [https://custom-scripts.sentinel-hub.com/sentinel-2/monthly_composite/#] is applied to overcome cloud cover issues by selecting the best available pixels of the last 30 days before the wildfire event.
- Sentinel-1 images are also cropped using the bounding box coordinates. Since cloud cover is not an issue with SAR images, there is no need to generate a mosaic. Thus, the most recent image before the event is selected. Sentinel-1 images are also cropped using the bounding box coordinates. Both ascending and descending images are provided.
- Burned area mask, a boolean mask of the burned area is also provided based on the rasterized EFFIS vector data on the Sentinel-2 grid.

The EO4WildFires dataset contains a rich set of features (Table 1) that can be used for wildfire severity prediction. The combination of meteorological data and satellite images can provide insights into the environmental conditions that can lead to the outbreak of wildfires. The wildfire event labels provided by the EFFIS system can be used to train predictive models for early detection and response to wildfire events.

The mosaicking process is used for creating a monthly composite image using Sentinel-2 satellite imagery. The process selects the best pixel for each date in the last 31 days based on a ratio of the available bands. The selection process is designed to avoid cloud cover, and the criteria used depends on the level of blue in the image. If blue is less than 0.12, the date with the maximum ratio of band B08 against band B02 is chosen. If no pixel is available above this threshold, the date with the maximum ratio of band B03 against band B02 is selected when blue is less than 0.45. If water is detected in the image, the date with the maximum ratio of band B02 against band B08 is chosen. If snow is detected, the median of the scene with snow is used. The resulting composite image provides a cloud-free representation of the last 31 days in the selected region.

A total of 31,730 wildfire events ranging from 2018 to 2022 across 45 countries, covering 8,707 level-4 administrative areas. We have used the GDAM (https://gadm.org/) database to correlate the polygons of the detected events from EFFIS with the level-4 administrative boundaries. After analysing the distribution of the data, it was found that the median size of a wildfire is 31.0 hectares while the mean is 128.77 hectares. An impressive 54,769.0 hectares was the largest recorded wildfire in 2021 in Antaya, Turkey, while the second largest wildfire was 51,245.0 hectares in 2021 in Evoia, Greece. The dataset is openly available at: 10.5281/zenodo.7762564

Figure 1 depicts on the map the level-4 administration boundaries of the recorded wildfire events (2018-2022) within the EO4WildFires dataset. For instance, we notice that specific areas in the Mediterranean have high concetrations in wildfire events. Figure 2 illustrates the total number of wildfires events for each year (2018-2022) per country. For instance, we may notice that the total number of detected events was higher in Ukraine in 2022 compared to previous years. Figure 3 depicts the median size of wildfires in hectares per country, where we can see that the Netherlands, Belgium and Austria demonstrate the largest values.

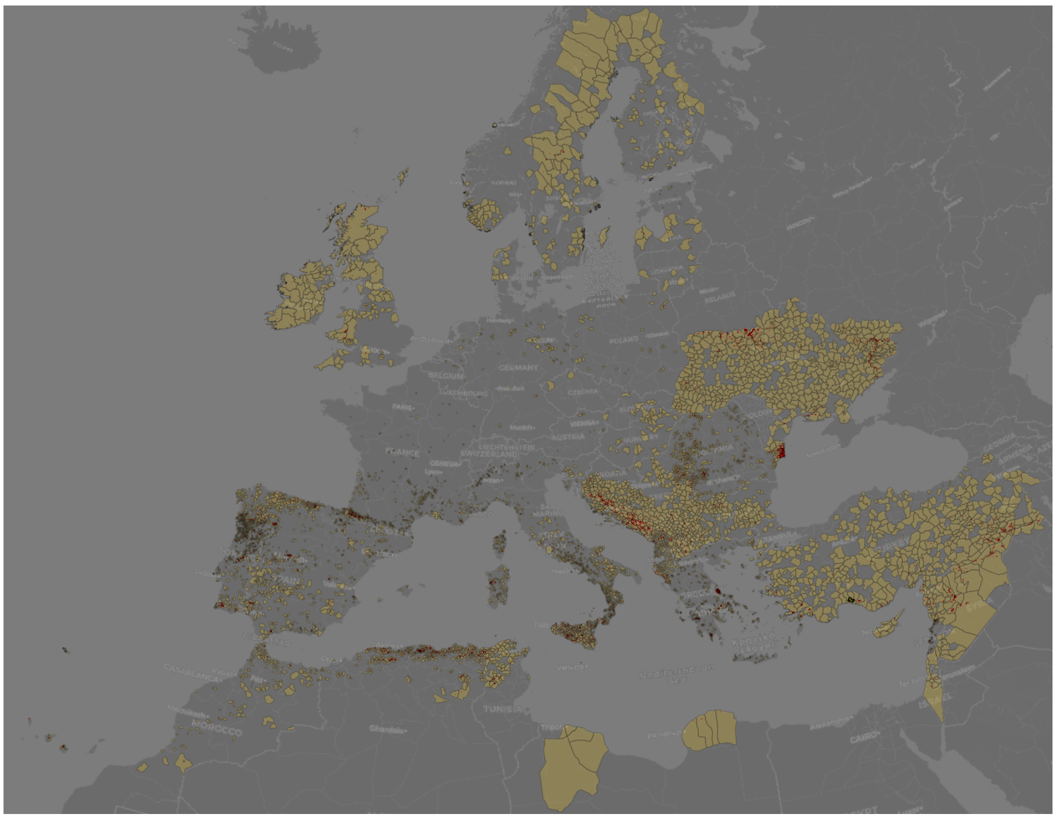

Figure 1. Wildfires events (2018-2022) with the corresponding affected level-4 administration boundaries – yellow polygons: administrative boundaries (level 4) of affected areas, red points: locations of wildfires events.

Figure 2. Total number of wildfires events for each year (2018-2022) per country. The size and color of the circles represent the total number of detected events.

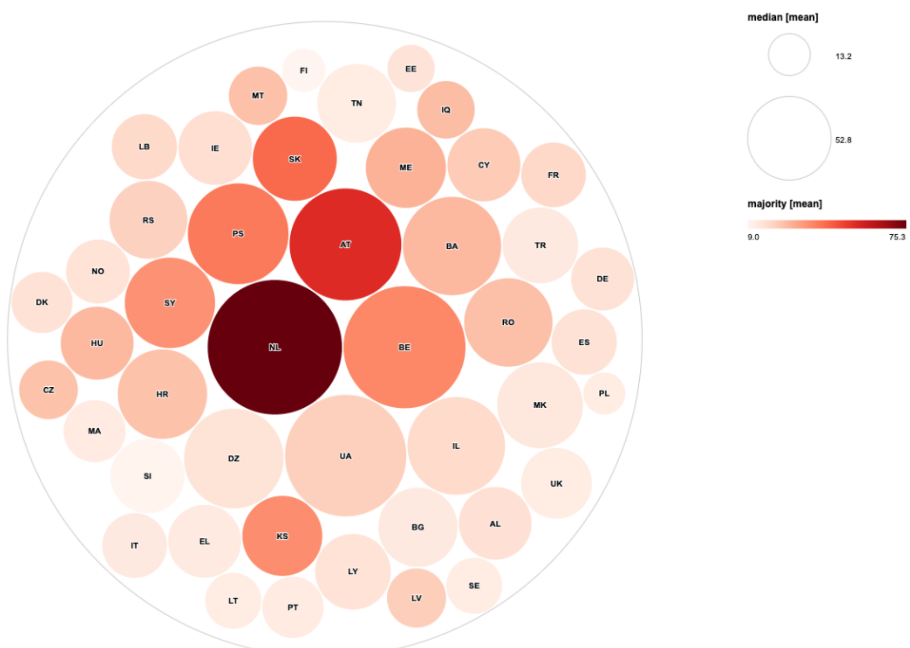

Figure 3. Median size of wildfires in hectares per country.

3. EXPERIMENTAL RESULTS FOR WILDFIRE SIZE PREDICTION

3.1 Experiments Description

This section describes three experiments aimed at predicting the severity (size) of a wildfire event utilizing the composed EO4WildFires dataset. In the first experiment, called **Meteo**, the input features used for prediction are only meteorological parameters listed in Table 1. In the second experiment, called **S1**, the input features used are the backscatter bands from Sentinel-1, and in the third experiment, called **S2**, the input features used are the mosaicked spectral bands from Sentinel-2. Specifically, for the S1 experiment, two different sub-experiments were conducted, one for the ascending and one for the descending orbits. This subdivision was put in place since the appearance of land features varies significantly depending on the illumination angle of the objects by the SAR beams. To maximize the potential of SAR data, ascending and descending data were isolated to minimize response variations due to illumination angles. This allows to focus the detection capabilities of the deep learning networks on the data variations that contribute into the prediction of the wildfire size.

The original EO4WildFires dataset is split into training, validation, and testing subsets, with 20,307 events in the training set, 5,077 events in the validation set, and 6,346 events in the test set. The division of the dataset is the same for all three experiments to ensure comparability of results. The goal of the experiments is to determine which input features provide the best prediction performance for the size of a wildfire event.

Heavily inspired by popular datasets like coco [10], three index files (train/val/test) are created that operate as literal file catalogues. Each row in the index files refers to a specific file in the disk.

3.2 Experiments Results

Sentinel-1 and Sentinel-2 can be illustrated as images that incorporate 3-dimension and 6-dimension channels, respectively. Therefore, it is logical to implement a popular feature extraction architecture to accumulate knowledge to a final classification layer. For this task, ResNet-32 [11] model was chosen, as it offers robust feature extraction given its total layer and parameter count. A modification was added to the optimizer, as it was trained using Lion [4] optimizer, which, given the latest advances, has shown great potential, as explained in Section 1.1.

Our experiments did not rely on pretrained weights for model initialization, but training was rather done from zero. The only difference between the experiments S1 and S2 was the input channel dimension in ResNet, as stated before. ResNet by default accepts 3 channels for 224x224 size images. S1 ascending and descending data match that description with minimum processing, but Sentinel 2A data do not. Finally, vanilla ResNet input dimensions, images and masks were padded to match integral multiples of 224 in each dimension, and later they were split into 224x224 chunks. The burned area, i.e. the classification label, was calculated as the sum of 1s in the padded image divided by (224*224).

Meteorological data are tabular and, thus, require minimum processing. Samples have been taken for thirty days prior to each event, therefore a 30-day sequence of 9 features (RH2M, T2M, PRECTOTCORR, WS2M, FRSNO, GWETROOT, SNODP, PRECSNOLAND, GWETTOP) can be directly fed into deep learning models primarily used for time series forecasting. A bidirectional Long Short-Term Memory (LSTM) network was chosen, consisting of 512 hidden dimension features and 3 layers. Outputs of the last hidden state of the model are propagated into a linear fully connected layer and then passed into loss and error functions, respectively.

For all training schedules, “Cross Entropy” loss was utilized, outputs of the model were passed into a Sigmoid function and then “Mean Absolute Error” was calculated to measure the distance between the predictions and the actually burned area. Training routine implements early stopping with minimum delta of 0.001 in validation loss and patience of 7 consecutive epochs before signaling trainer to stop. Finally, all models’ optimizers have a learning rate scheduler attached to them, namely “Reduce LROn Plateau” is utilized, which updates learning rate by 0.1 every 3 epochs.

The results of the three experiments, in terms of test loss and test error, are summarized in Table 2 and further discussed in Section 4.

Table 2. Experiment results.

	Test Loss (Cross Entropy)	**Test Error (MAE)**
Meteo	0.422	0.129

S1 Ascending	0.524	0.066
S1 Descending	0.517	0.062
S2	0.371	0.056

4. CONCLUSIONS

This paper presents EO4WildFires, a comprehensive benchmark dataset that integrates multispectral, SAR, and meteorological data to provide insights into the factors affecting wildfire severity. Through a series of experiments, we demonstrate the potential of using machine learning models to predict wildfire size based on various combinations of input data.

The EO4WildFires dataset addresses the scientific problem of forecasting the severity of future wildfire events, based on recent past meteorological (30-days before the event) data, and most recent past multispectral, and SAR images (earliest available images before the event). The dataset combines data from the EFFIS, Copernicus Sentinel-1 & 2, and NASA Power to create a comprehensive source of information for developing predictive models. The dataset, which includes 31,740 files covering 31,730 wildfire events across 45 countries, provides a rich set of features that can be used to train machine learning models to predict wildfire severity.

Three experiments were conducted to predict the size of a wildfire event using different input features: meteorological parameters (Meteo), backscatter bands from Sentinel-1 (S1), and mosaicked spectral bands from Sentinel-2 (S2). The ascending and descending orbits of Sentinel-1 were treated separately. The dataset was divided into training, validation, and testing subsets, with the same division across all experiments to ensure comparability of results. ResNet-32 was chosen as the feature extraction architecture for Sentinel-1 and Sentinel-2 data, while a bidirectional LSTM network was used for meteorological data. The goal of these experiments was to determine which input features provide the best prediction performance for wildfire size.

As shown in Table 2, the experiment with the lowest error was S2. Features from S2 represent the actual status (vegetation and soil moisture, fuel and flammability) of the forest, which suggests that such data, when used as input, contribute into the prediction of wildfire severity. On the other hand, the Meteo experiment has the highest error. Features from Meteo include only the atmospheric conditions, which have don't capture the full extent of the status of different vegetation conditions. The SAR experiments present an reasonably low error, which can be attributed to the fact that although SAR data do not directly measure the vegetation status, they do provide information about the vegetation canopy and forest internal structure. In future work, the authors will continue to experiment with further combinations of data inputs and different network architectures.

By leveraging this dataset, researchers, first responders and policymakers can be better informed by accurate and up-to-date information on wildfire events and their potential severity.

ACKNOWLEDGEMENT

This study has been conducted in the framework of SILVANUS project. This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101037247. The contents of this publication are the sole responsibility of the authors and can in no way be taken to reflect the views of the European Commission.

REFERENCES

[1] Jiahui Yu, Zirui Wang, Vijay Vasudevan, Legg Yeung, Mojtaba Seyedhosseini, Yonghui Wu, "CoCa: Contrastive Captioners are Image-Text Foundation Models", Computer Vision and Pattern Recognition, (2022).

[2] Ze Liu, Han Hu, Yutong Lin, Zhuliang Yao, Zhenda Xie, Yixuan Wei, Jia Ning, Yue Cao, Zheng Zhang, Li Dong, Furu Wei, Baining Guo, " Swin Transformer V2: Scaling Up Capacity and Resolution," Computer Vision and Pattern Recognition (2022).

[3] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, Neil Houlsby, " An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," Computer Vision and Pattern Recognition, (2021).

[4] Xiangning Chen, Chen Liang, Da Huang, Esteban Real, Kaiyuan Wang, Yao Liu, Hieu Pham, Xuanyi Dong, Thang Luong, Cho-Jui Hsieh, Yifeng Lu, Quoc V. Le, "Symbolic Discovery of Optimization Algorithms," Machine Learning (2023).

[5] Team, M. F. (2022, November 18). MODIS Active Fire and Burned Area Products - Home. Retrieved from https://modis-fire.umd.edu.

[6] Ying, L., Shen, Z., Yang, M., and Piao, S., "Wildfire Detection Probability of MODIS Fire Products under the Constraint of Environmental Factors: A Study Based on Confirmed Ground Wildfire Records," Remote Sensing 11(24), 3031 (2019).

[7] Younes Oulad Sayad, Hajar Mousannif, Hassan Al Moatassime, "Predictive modeling of wildfires: A new dataset and machine learning approach," Fire Safety Journal (2019)

[8] F. Huot, R. L. Hu, N. Goyal, T. Sankar, M. Ihme and Y. -F. Chen, "Next Day Wildfire Spread: A Machine Learning Dataset to Predict Wildfire Spreading From Remote-Sensing Data," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-13, (2022).

[9] Suwei Yang, Massimo Lupascu, Kuldeep S. Meel, "Predicting Forest Fire Using Remote Sensing Data And Machine Learning," The Thirty-Fifth AAAI Conference on Artificial Intelligence (2021).

[10] Tsung-Yi Lin et al., 2014. Microsoft COCO: Common Objects in Context. *CoRR*, abs/1405.0312. Available at: http://arxiv.org/abs/1405.0312.

[11] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.