

# Predicting Next-Day Wildfire Spread with Time Series and Attention

Saad Lahrichi  
University of Missouri  
Columbia, MO 65201

saad.lahrichi@missouri.edu

Jesse Johnson  
University of Montana  
Missoula, MT 59802

jesse.johnson@umt.edu

Jordan Malof  
University of Missouri  
Columbia, MO 65201

jmdrp@missouri.edu

## Abstract

Recent research has demonstrated the potential of deep neural networks (DNNs) to accurately predict next-day wildfire spread, based upon the current extent of a fire and geospatial rasters of influential environmental covariates e.g., vegetation, topography, climate, and weather. In this work, we investigate a recent transformer-based model, termed the SwinUnet, for next-day wildfire prediction. We benchmark Swin-based models against several current state-of-the-art models on WildfireSpreadTS (WFTS), a large public benchmark dataset of historical wildfire events. We consider two next-day fire prediction scenarios: when the model is given input of (i) a single previous day of data, or (ii) five previous days of data. We find that, with the proper modifications, SwinUnet achieves state-of-the-art accuracy on next-day prediction for both the single-day and multi-day scenarios. SwinUnet’s success depends heavily upon utilizing pre-trained weights from ImageNet. Consistent with prior work, we also found that models with multi-day-input always outperformed models with single-day input.

## 1. Introduction

Wildfires are a global cause of concern that have severe human, economical, and environmental impacts, with the average annual economic burden from wildfires falling between \$71.1 billion and \$347.8 billion [29]. Due to climate change, many regions are getting hotter and drier, which lengthens wildfire seasons and exacerbates their effect. In order to better manage, mitigate, and prevent wildfires, accurately predicting their spread is essential. In this work, we focus on the problem of next-day wildfire spread prediction, where we are provided with current and/or historical information about a particular wildfire, and then tasked with predicting its spatial extent on the following day.

A variety of approaches have been investigated to solve this problem, such as those based upon machine learning models [16, 3, 7], or physics-based and observationally in-

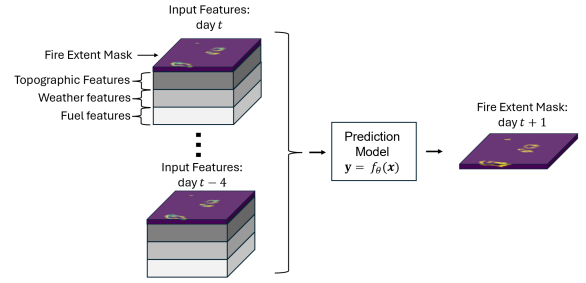


Figure 1. The wildfire prediction models take as input a geospatial raster of several variables: fuel indicators, topography, and weather features, alongside the current day fire mask. We consider two prediction scenarios: one in which the model receives input features only for the current day, denoted  $t$ , and another in which we provide the model five previous days of features as input. In either case, the model must predict a binary mask indicating the extent of the fire on the following day,  $t+1$ .

formed models [8, 1, 9]. In this work, however, we focus on a promising emerging class of techniques that utilize high-capacity machine learning models - namely deep neural networks (DNNs) - to predict wildfire spread based upon high-dimensional explanatory input data. These input data typically comprise a geospatial map of the current extent of the fire, as well as maps of explanatory features such as topography, climate, weather, and vegetation indices. Based upon these input data, the model is tasked with producing a geospatial map, or an image, reflecting the spatial extent of the fire on the following day, i.e. the spread of fire into regions that have not burned, and the extinction of fire in regions that have ceased burning. See Fig. 1 for an illustration.

DNN-based models have achieved impressive prediction accuracy, and have garnered substantial attention in recent years for wildfire modeling, including specifically next-day wildfire prediction. A variety of DNN-based models have been proposed to solve this problem, including convolutional models [20, 4, 22], attention-based models such as transformers [28], and spatio-temporal models [23, 2]. Most existing research has focused on next-day prediction

where only explanatory data from the current day is provided (day  $t$  only, in Fig. 1). However, recent research found that models utilizing a time-series of historical input data can achieve greater prediction accuracy [12], suggesting this as an important new direction in next-day wildfire prediction.

**Contributions of this Work** In this work, we investigate attention-based models for time-series (i.e., multi-day input) next-day wildfire prediction. The wildfire spread prediction task is essentially a binary segmentation task (burning, not burning), and therefore we focus our investigation on a recent model, termed the SwinUnet [5] which has been found to be highly effective for solving such tasks in a variety of contexts. We adapt the SwinUnet for the wildfire prediction and evaluate its performance on the recent WildFire-SpreadTS (WFTS) benchmark [12]. We chose the WFTS benchmark because it is a public benchmark designed to support time-series wildfire prediction, and which employs a rigorous 12-fold leave-one-year-out cross-validation with a large collection of historical wildfire events. We compare our SwinUnet model to the reported performance of current state-of-the-art models on the WFTS benchmark.

Our results indicate that the SwinUnet performs poorly when trained from scratch. However, when we initialize training with pre-trained weights (e.g., from ImageNet), the SwinUnet model achieves a new state-of-the-art accuracy on the WFTS benchmark, both for the single-day and time-series input data. Although other models benefit from pre-trained weights, the SwinUnet obtains the greatest benefits. Corroborating recent research [12], we also find that models trained on time-series always outperform their single-day counterparts.

The rest of the paper is structured as follows: Section 2 reviews related works, Section 3 describe our adopted benchmark and methods, Section 4 details our experimental setting and results, and Section 5 summarizes our findings.

## 2. Related Works

**Next-day wildfire prediction** To apply DNNs to fire spread prediction, recent research has framed the spread problem as a semantic segmentation one, and developed multiple datasets to support this framing. [15] created Next-Day Wildfire Spread, a dataset for mono-temporal fire spread prediction. They developed a custom convolutional autoencoder that takes as input various explanatory variables and outputs a binary mask indicating fire presence at each pixel. Similarly, [12] extended the dataset to include multi-temporal prediction, added more explanatory variables, and higher resolution fire masks. They achieved the best performance using a standard Unet model and a Unet model with temporal attention (UTAE) [11]. Other datasets aim to predict beyond next-day and forecast fire behavior several days in advance. For example, [25] de-

veloped SeasFire Cube and trained Unet++ models [32] for medium-term fire prediction, between 8 and 64 days. [17] improved upon the collected data cube and found that the LSTM and ConvLSTM models outperformed the Fire Weather Index (FWI). In FireSight, [14] collected a dataset using remote sensing data from 20 datasets, and trained a 3D UNet model to model short-term fire hazard, between 3 and 8 days.

**Other approaches** Aside from segmentation, other formulations have been developed for modeling fire spread using DL. [27] used a CNN-based Reinforcement Learning (RL) model that predicts the best action as burn or no burn given current conditions. [13] developed a probabilistic cellular automata model to simulate wildfire spread. In [19], the authors developed Sim2Real-Fire, a synthetic, high-resolution dataset for fire spread forecast and backtracking and outperformed the considered baselines using a custom Transformer model. We refer the readers to [31] for a more comprehensive review of DL for wildfire prediction.

**Next-day wildfire prediction with time-series** In contrast to most existing work (e.g., [18, 10, 28, 30]), our work here focuses on utilizing a time-series of features for next-day wildfire spread prediction, which has been cited as an important emerging area of research [10, 12, 18]. Historically, time-series modeling has been challenging due to the lack of appropriate public datasets to train and evaluate models for this task. Recently, [12], building upon the work of [15], developed the first multi-temporal dataset for time-series prediction. Crucially, they found that models using a time-series of input tend to be more accurate than those using a single day, making this an important emerging area of research. We build on this existing time-series prediction research by investigating attention-based models: specifically the recent SwinUnet model [21], which has been found extremely effective in a variety of contexts for segmentation tasks.

## 3. Materials and Methods

**The WFTS Benchmark** In this work, we employ the WildFireSpreadTS (WFTS) dataset [12]. The dataset includes 607 wildfire events across the western United States between 2018 and 2021, totaling 13,607 daily multi-channel images. These 23 channels include data on active fires, weather, topography, and vegetation, resampled to a common resolution of 375 meters, providing a rich multi-modal and multi-temporal framework for modeling fire spread. A key feature of this benchmark is a rigorous 12-fold cross-validation model evaluation procedure. Each fold of the cross-validation includes all wildfire events from a single year, so that the trained models are always evaluated on wildfire events from a previously unseen year, reflecting real-world use of wildfire prediction models.

**SwinUnet** SwinUnet [5] is a pure transformer-based Unet-shaped model that was first proposed for medical imagery segmentation. The model replaces the convolution blocks of the Unet with Swin Transformer blocks [21], including them throughout the encoder, bottleneck, and decoder. They also rely on patch merging and patch expansion layers in the encoder and decoder, respectively, to downsample the input features and then upsample the extracted features and produce the segmentation mask. Finally, they preserve skip connections to concatenate shallow and deep features. The SwinUnet outperformed the Unet [26], ViT [6], Att-Unet [24], and TransUnet [6] on two medical benchmark datasets. Its state-of-the-art performance, ability to learn both global and long-range dependencies, and use of the more efficient Swin blocks make it a good candidate for our task. Since the model was developed for RGB images, we modify the `in_chans` parameter to take in the number of channels of our multi-modal inputs instead of 3.

**Model pre-training** To evaluate the effect of pre-training on the SwinUnet model, we load the `swin-tiny-patch4-window7-224` weights from HuggingFace onto each of our Swin blocks. These weights correspond to a Swin Transformer trained on ImageNet at 224x224 resolution. We zero-pad our input images (128x128) to match the expected input dimensions and benefit from the pre-trained weights. As for the Unet models, we follow [12] and use the `segmentation_models.pytorch` implementation, and set `encoder_weights` to `imagenet`, which loads a model with ImageNet pre-trained weights.

## 4. Experiments

To assess the benefits of using transformer-based models, we follow the original WildfireTS experimental design [12]. We train the SwinUnet model on two years of input, and validate and test on a year each. Since the dataset includes four years of input, this results in a twelve fold cross-validation. We report the average test average precision (AP) scores across all twelve folds as well as the standard deviation. Our main results can be found in Table 1, where we compare to the best-performing convolution-based model (Res18-Unet) and the best-performing model overall (UTAE). Similarly to the Unet models, the SwinUnet cannot handle multi-temporal data natively. Therefore, we follow the processing used for the Unet models and simply concatenate the features for all time steps along the channel dimension. We also train all models ourselves, and replicate the performance reported in the original paper.

### 4.1. Results

**SwinUnet Achieves the Highest Accuracy** Our experimental results are reported in Table 1, where the results are stratified by whether the models utilized a single day of input, or a time-series of five days of input. We observe that in both cases the SwinUnet achieves the highest prediction accuracy of all models considered, confirming the suitability of attention-based models for the wildfire spread prediction task.

**Single-day vs Time-series Prediction** Our results still indicate that the best overall accuracy is achieved with time-series input, as opposed to single-day input. This is consistent with recent findings in [12], but demonstrates that the pattern continues to hold even with the introduction of a new model with improved overall accuracy. This suggests that multiple days of historical data about a wildfire may allow the models to infer key properties of a wildfire that are unavailable in the single day. This generally suggests that time-series wildfire prediction is a promising general direction for next-day wildfire prediction.

**Pre-training impact** We evaluate whether initializing our models with pre-trained weights is beneficial. In Table 1, we show the results on the Vegetation features when loading our models with PASTIS pre-trained weights for the UTAE [11] and ImageNet pre-trained weights for the other models. While pre-training generally has a modest impact on the performance of most models, it greatly benefits the SwinUnet model. The average test AP of the SwinUnet improves from 0.351 when training from scratch to 0.383 when using pre-training using 1-day inputs, and from 0.365 to 0.404 using 5-day inputs. Crucially, these results show that the SwinUnet did not outperform other models simply because of its pre-training, and thereby support the conclusion that attention-based models, such as SwinUnet, are better suited to the wildfire classification task.

**Model size impact** In order to account for the number of parameters used by each model, we also include a Res50-Unet to the models benchmarked by [12]. The results in Table 1 suggest that even with more parameters, the Res50-Unet underperforms when compared to the SwinUnet, achieving lower Test AP across both from scratch and pre-training settings, and both single and multi-day inputs. This further proves the potential of Transformer models for the wildfire spread prediction task, and their higher predictive accuracy.

**Qualitative Evaluation of Performance.** Fig. 2 presents four example predictions made by the SwinUnet. In many cases, the model successfully predicts general wildfire progression, and most error is caused by error in the fine-scale detail of the fire. For example, in the top two rows of Fig. 2 the model correctly predicts that the fires will shrink overall, but does not predict the exact shape of the fire after its recession. By contrast, we find that the

Table 1. Mean test AP  $\pm$  standard deviation using vegetation features only when training with 1 and 5 input days. SwinUnet with spatial attention and ImageNet-pretraining sets a new state-of-the-art performance on this benchmark. Cited rows represent accuracy reported on our benchmark from previous publications. Results style: **best**, second best. \* indicates a pre-trained model on ImageNet; \*\* pre-trained on PASTIS [11]

Model	Input days	Test AP	# Params
Res50-Unet*	1	$0.320 \pm 0.078$	32.6M
Res50-Unet	1	$0.321 \pm 0.084$	32.6M
SwinUnet*	1	<b><math>0.383 \pm 0.087</math></b>	27.2M
SwinUnet	1	$0.351 \pm 0.087$	27.2M
Res18-Unet*	1	$0.314 \pm 0.080$	14.3M
Res18-Unet[12]	1	$0.328 \pm 0.090$	14.3M
Res50-Unet*	5	$0.329 \pm 0.071$	32.9M
Res50-Unet	5	$0.330 \pm 0.083$	32.6M
SwinUnet*	5	<b><math>0.404 \pm 0.081</math></b>	27.3M
SwinUnet	5	$0.365 \pm 0.085$	27.2M
Res18-Unet*	5	$0.331 \pm 0.075$	14.4M
Res18-Unet	5	$0.335 \pm 0.070$	14.4M
UTAE**	5	$0.381 \pm 0.103$	1.1M
UTAE [12]	5	$0.372 \pm 0.088$	1.1M

model often struggles with events where there are relatively few fire pixels on the current day. In the last two rows of Fig. 2 for example, the model fails to identify the suddenly large growth of the fire. In other words, the model performs poorly when predicting ignition or onset, but improves significantly on tasks related to established fire progression or spread.

## 5. Conclusion

In this work, we evaluated the benefits of using transformer-based models for wildfire spread prediction on the WildfireTS dataset. Using a SwinUnet model with ImageNet pre-trained Swin blocks significantly outperformed fully convolutional baselines, establishing a new state-of-the-art benchmark performance. These results highlight the suitability of transformers for this task, particularly in the multi-temporal setting.

Future work should address the challenges of labeled data scarcity and domain shift by using Self-Supervised Learning (SSL) to train transformers, which are designed to handle missing input features. Another important area of future research is the development of models that integrate both spatial and temporal attention.

## References

[1] Martin E Alexander and Miguel G Cruz. Evaluating a model for predicting active crown fire rate of spread using wildfire observations. *Canadian Journal of Forest Research*, 36(11):3015–3028, 2006. 1

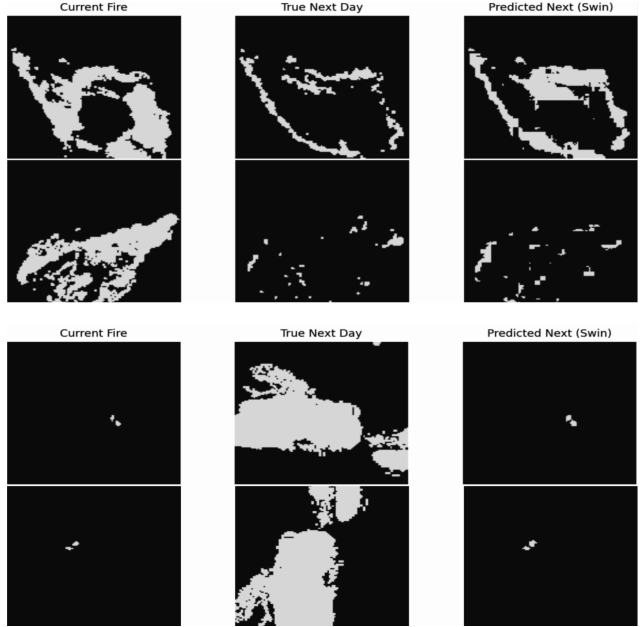


Figure 2. Example predictions made by the SwinUnet model. The leftmost column represents the current day, the middle column represents the ground truth next day, and the rightmost column represents the prediction. The first two rows represent examples where the model achieved relatively high accuracy, while the second two rows are examples of cases with poor accuracy.

[2] Andrew Bolt, Carolyn Huston, Petra Kuhnert, Joel Janek Dabrowski, James Hilton, and Conrad Sanderson. A spatio-temporal neural network forecasting approach for emulation of firefront models. In *2022 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, pages 110–115. IEEE, 2022. 1

[3] Karol Bot and José G Borges. A systematic review of applications of machine learning techniques for wildfire management decision support. *Inventions*, 7(1):15, 2022. 1

[4] John Burge, Matthew R Bonanni, R Lily Hu, and Matthias Ihme. Recurrent convolutional deep neural networks for modeling time-resolved wildfire spread behavior. *Fire Technology*, 59(6):3327–3354, 2023. 1

[5] Hu Cao, Yueyue Wang, Joy Chen, Dongsheng Jiang, Xi-aopeng Zhang, Qi Tian, and Manning Wang. Swin-unet: Unet-like pure transformer for medical image segmentation. In *European conference on computer vision*, pages 205–218. Springer, 2022. 2, 3

[6] Jieneng Chen, Yongyi Lu, Qihang Yu, Xiangde Luo, Ehsan Adeli, Yan Wang, Le Lu, Alan L Yuille, and Yuyin Zhou. Transunet: Transformers make strong encoders for medical image segmentation. *arXiv preprint arXiv:2102.04306*, 2021. 3

[7] Khaled Chetehouna, Eddy El Tabach, Loubna Bouazaoui, and Nicolas Gascoin. Predicting the flame characteristics and rate of spread in fires propagating in a bed of pinus pinaster using artificial neural networks. *Process Safety and Environmental Protection*, 98:50–56, 2015. 1



- [8] Mark A. Finney. *FARSITE: Fire Area Simulator-model development and evaluation*. 1998. [1](#)
- [9] Mark A Finney. An overview of flammmap fire modeling capabilities. In *In: Andrews, Patricia L.; Butler, Bret W., comps. 2006. Fuels Management-How to Measure Success: Conference Proceedings. 28-30 March 2006; Portland, OR. Proceedings RMRS-P-41. Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station. p. 213-220, volume 41, 2006.* [1](#)
- [10] Jack Fitzgerald, Ethan Seefried, James E Yost, Sangmi Pallickara, and Nathaniel Blanchard. Paying attention to wildfire: Using u-net with attention blocks on multimodal data for next day prediction. In *Proceedings of the 25th International Conference on Multimodal Interaction*, pages 470–480, 2023. [2](#)
- [11] Vivien Sainte Fare Garnot and Loic Landrieu. Panoptic segmentation of satellite image time series with convolutional temporal attention networks. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4872–4881, 2021. [2](#), [3](#), [4](#)
- [12] Sebastian Gerard, Yu Zhao, and Josephine Sullivan. Wildfirespreadts: A dataset of multi-modal time series for wildfire spread prediction. *Advances in Neural Information Processing Systems*, 36:74515–74529, 2023. [2](#), [3](#), [4](#)
- [13] Rohit Ghosh, Jishnu Adhikary, and Rezki Chemlal. Fire spread modeling using probabilistic cellular automata. In *Asian Symposium on Cellular Automata Technology*, pages 45–55. Springer, 2024. [2](#)
- [14] Julia Gottfriedsen, Johanna Strebl, Max Berrendorf, Martin Langer, and Volker Tresp. Firesight: Short-term fire hazard prediction based on active fire remote sensing data. [2](#)
- [15] Fantine Huot, R Lily Hu, Nita Goyal, Tharun Sankar, Matthias Ihme, and Yi-Fan Chen. Next day wildfire spread: A machine learning dataset to predict wildfire spreading from remote-sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–13, 2022. [2](#)
- [16] Sadegh Khanmohammadi, Mehrdad Arashpour, Emadaldin Mohammadi Golafshani, Miguel G Cruz, Abbas Rajabifard, and Yu Bai. Prediction of wildfire rate of spread in grasslands using machine learning methods. *Environmental Modelling & Software*, 156:105507, 2022. [1](#)
- [17] Spyros Kondylatos, Ioannis Prapas, Michele Ronco, Ioannis Papoutsis, Gustau Camps-Valls, María Piles, Miguel-Ángel Fernández-Torres, and Nuno Carvalhais. Wildfire danger prediction and understanding with deep learning. *Geophysical Research Letters*, 49(17):e2022GL099368, 2022. [2](#)
- [18] Bronte Sihan Li and Ryan Rad. Wildfire spread prediction in north america using satellite imagery and vision transformer. In *2024 IEEE Conference on Artificial Intelligence (CAI)*, pages 1536–1541. IEEE, 2024. [2](#)
- [19] Yanzhi Li, Keqiu Li, LI GUOHUI, Chanqing Ji, Lubo Wang, Die Zuo, Qing Guo, Feng Zhang, Manyu Wang, Di Lin, et al. Sim2real-fire: A multi-modal simulation dataset for forecast and backtracking of real-world forest fire. In *The Thirty-eight Conference on Neural Information Processing Systems Datasets and Benchmarks Track*. [2](#)
- [20] Shuwen Liu, Lin Cao, Chuanying Lin, Yuxuan Dai, Xingdong Li, Sanping Li, Shufa Sun, and Dandan Li. Fire spread prediction model based on multi-scale convolutional neural network. *Multimedia Tools and Applications*, pages 1–22, 2024. [1](#)
- [21] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 10012–10022, 2021. [2](#), [3](#)
- [22] Mohammad Marjani, Masoud Mahdianpari, and Fariba Mohammadimanesh. Cnn-bilstm: A novel deep learning model for near-real-time daily wildfire spread prediction. *Remote Sensing*, 16(8):1467, 2024. [1](#)
- [23] Dimitrios Michail, Lefki-Ioanna Panagiotou, Charalampos Davalas, Ioannis Prapas, Spyros Kondylatos, Nikolaos Ioannis Bountos, and Ioannis Papoutsis. Seasonal fire prediction using spatio-temporal deep neural networks. *arXiv preprint arXiv:2404.06437*, 2024. [1](#)
- [24] Ozan Oktay. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*, 2018. [3](#)
- [25] Ioannis Prapas, Akanksha Ahuja, Spyros Kondylatos, Ilektra Karasante, Eleanna Panagiotou, Lazaro Alonso, Charalampos Davalas, Dimitrios Michail, Nuno Carvalhais, and Ioannis Papoutsis. Deep learning for global wildfire forecasting. *arXiv preprint arXiv:2211.00534*, 2022. [2](#)
- [26] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pages 234–241. Springer, 2015. [3](#)
- [27] William L Ross. Being the fire: A cnn-based reinforcement learning method to learn how fires behave beyond the limits of physics-based empirical models. In *NeurIPS 2021 Workshop on Tackling Climate Change with Machine Learning*, 2021. [2](#)
- [28] Kamen Shah and Maria Pantoja. Wildfire spread prediction using attention mechanisms in u-net. In *2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, pages 1–6. IEEE, 2023. [1](#), [2](#)
- [29] Douglas Thomas, David Butry, Stanley Gilbert, David Webb, Juan Fung, et al. The costs and losses of wildfires. *NIST special publication*, 1215(11):1–72, 2017. [1](#)
- [30] Hongtao Xiao, Yingfang Zhu, Yurong Sun, Gui Zhang, and Zhiwei Gong. Wildfire spread prediction using attention mechanisms in u2-net. *Forests*, 15(10):1711, 2024. [2](#)
- [31] Zhengsen Xu, Jonathan Li, Sibbo Cheng, Xue Rui, Yu Zhao, Hongjie He, and Linlin Xu. Wildfire risk prediction: A review. *arXiv preprint arXiv:2405.01607*, 2024. [2](#)
- [32] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4*, pages 3–11. Springer, 2018. [2](#)