CS230

Wildfire Prediction Using Deep Learning Models for Remote Sensing Data

Project Category: Computer Vision

Raghav Sharma
Department of Civil and Environmental Engineering
Stanford University
raghavsh@stanford.edu
SUNet ID: raghavsh

Rishabh Aggarwal
Graduate School of Business
Stanford University
arishabh@stanford.edu
SUNet ID: arishabh

Abstract

The abstract should consist of 1 paragraph describing the motivation for your paper and a high-level explanation of the methodology you used/results obtained.

1 Introduction

Climate change is making extreme weather events more probable and intense and it is becoming increasingly important to estimate extreme event risk to build informed resilience. Wildfires are one such event of extreme importance, especially in California where wildfire risk is high. For authorities to implement consequential wildfire management, accurate estimation of wildfire likelihood is essential.

We propose a deep learning model for predicting wildfire occurrence based on remote sensing data of factors influencing wildfires such as topography, max/min temperature, precipitation, drought index, wind speed, Normalized Difference Vegetation Index (NDVI), and humidity. Given the features influencing wildfire, we predict the spread of wildfire the *next* day.

2 Related Work

The main source of remote-sensing data is Google Earth Engine which is also used by Huot et al. [2020, 2021]. As a result, Huot et al. [2021] is the closest paper and forms the basis of our analysis. In addition, Prapas et al. [2021] selects learning algorithms chosen carefully based on the type of dataset used (spatial vs temporal) and serves as a guide to model architecture. Similarly, Sayad et al. [2019] evaluate their models via various metrics whereas Ghorbanzadeh et al. [2019], Pham et al. [2020] provide validation and performance evaluation techniques for different models. Further, Jain et al. [2020] provides an exhaustive review of the machine learning applications in wildfire science and Barmpoutis et al. [2020] reviews the literature on early wildfire detection using remote sensing.

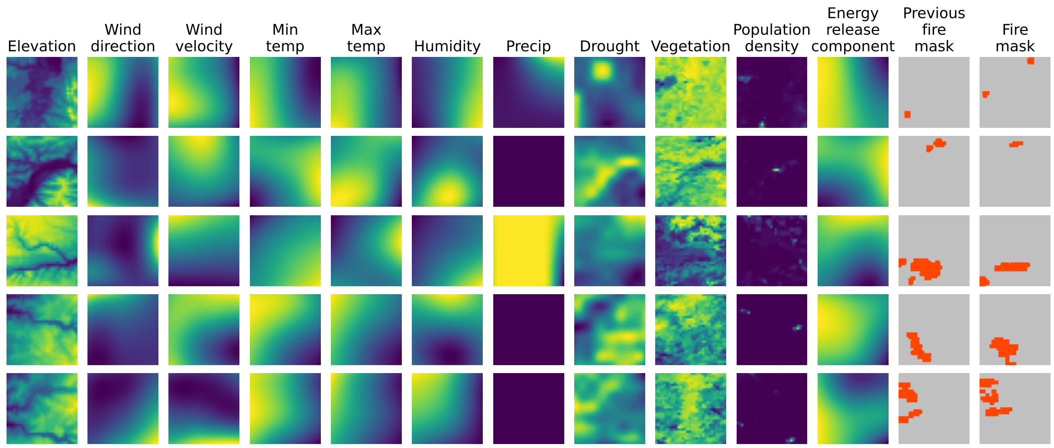

Figure 1: Examples from the dataset

3 Dataset and Features

We use the 'Next Day Wildfire Spread' data set compiled by Huot et al. [2021] from various data sources available through Google Earth Engine (GEE).[1] Here, we provide a brief description of the dataset and the features.[2]

The dataset combines historical wildfire events with remote sensing images from GEE across United States from 2012 to 2020. The data is extracted as images of 64km $\times$ 64km regions at 1 km resolution and includes input features that influence wildfire. There are 11 input features: elevation, wind direction and wind speed, minimum and maximum temperatures, humidity, precipitation, drought index, normalized difference vegetation index (NDVI), population density and energy release component (ERC). Although the data sources, that the input features are extracted from, have different spatial resolutions, all the data is aligned to 1 km resolution, which corresponds to the spatial resolution of the fire masks.

The historical wildfire dataset is processed to represent the fire information as a fire mask over each 64km $\times$ 64km region, showing the locations of 'fire' versus 'no fire', with an extra classification for uncertain labels and includes both the fire mask at time t, denoted as 'previous fire mask' and at time $t+1$ day, denoted as 'fire mask' to provide two snapshots of the fire spreading pattern. Fires separated by more than 10 km are considered to be belonging to a different fire. For the purpose of machine learning algorithm, the fire mask at time t is considered as an input feature and the 'fire mask' at time $t+1$ day as labels. For characterizing fire spreading, only the samples for which the 'previous fire mask' contains any fire at all are kept. Also, we drop the samples containing 'uncertain' labels.

The objective is to predict where the fire will spread on date $t+1$ given input features at date t.

3.1 Data Preprocessing

There are three main pre-processing elements added to the dataset. First, the input features, except the fire masks, are clipped at 0.1% and 99.9% percentiles for each feature. Second, each feature is then normalized by subtracting the mean and dividing by the standard deviation. These statistics are computed over the training set after clipping. Third, the dataset is augmented by randomly cropping 32 km $\times$ 32 km from the original 64 km $\times$ 64 km regions. The final dataset consists of 13,602 examples. The dataset is split between training, development and test sets according to 8:1:1 ratio.

Figure 1 displays examples from the dataset where each row represent one example of 32 km $\times$ 32 km at 1 km resolution. Each row corresponds to the 11 input features, previous fire mask at time t at a particular location and the fire mask at time $t+1$. As mentioned, the fire mask at date t is considered

[1] The dataset is made public on Kaggle by the authors.

[2] Further details regarding data sources and aggregation can be found in Huot et al. [2020, 2021].

as an input feature and labeled as the 'previous fire mask' whereas 'fire mask' corresponds to the $t+1$. In the fire masks, red denotes fire, while grey implies no fire.

4 Model Architecture

Since the dataset consists of spatial features, we employ a U-Net convolutional neural net architecture. A U-Net CNN frames the prediction task as an image segmentation problem where we classify each area as either containing fire or no fire given the location of the fire on the previous day and other input features.

Figure 2 shows the architecture for the image segmentation problem. The input to the network is a $32 \times 32 \times 12$ image and the output is a $32 \times 32 \times 1$ image of fire masks on day $t+1$. All convolutionals are 3×3 with a stride of 1 and same padding, max pooling is 2×2 and we use ReLU activation function for the hidden layers and a sigmoid activation for the output layer. A U-Net CNN model architecture shrinks the image size in the contractionary paths and expands it back to the output size during the expansion path. We use 4 down Conv blocks with 16, 32, 64 and 128 filters where each Conv block is followed by a max pooling. Each Conv block consists of a 3×3 Conv2D block, dropout, another 3×3 Conv2D block followed by batch normalization and ReLU activation (shown at bottom of Figure 2). During the expansion path, we use 2×2 upsample and same Conv (up) block. The last layer has 1 filter with 1×1 Conv2D block and sigmoid activation function for the predicted fire mask image.

4.1 Training Details and Hyperparameters tuning

5 Evaluation

We will evaluate the results using three main quantitative metrics - Precision, Recall, and the Area under the Receiver Operator Curve. Qualitatively, the models will produce predictive images of the fire masks which can then be compared to the actual images from Google Earth Engine.

6 GitHub Link

We have posted the code to pre-process and visualize the data on Github.

References

Panagiotis Barmpoutis, Periklis Papaioannou, Kosmas Dimitropoulos, and Nikos Grammalidis. A review on early forest fire detection systems using optical remote sensing. *Sensors*, 20(22):6442, 2020.

Omid Ghorbanzadeh, Khalil Valizadeh Kamran, Thomas Blaschke, Jagannath Aryal, Amin Naboureh, Jamshid Einali, and Jinhu Bian. Spatial prediction of wildfire susceptibility using field survey gps data and machine learning approaches. *Fire*, 2(3):43, 2019.

Fantine Huot, R Lily Hu, Matthias Ihme, Qing Wang, John Burge, Tianjian Lu, Jason Hickey, Yi-Fan Chen, and John Anderson. Deep learning models for predicting wildfires from historical remote-sensing data. *arXiv preprint arXiv:2010.07445*, 2020.

Fantine Huot, R Lily Hu, Nita Goyal, Tharun Sankar, Matthias Ihme, and Yi-Fan Chen. Next day wildfire spread: A machine learning data set to predict wildfire spreading from remote-sensing data. *arXiv preprint arXiv:2112.02447*, 2021.

Piyush Jain, Sean CP Coogan, Sriram Ganapathi Subramanian, Mark Crowley, Steve Taylor, and Mike D Flannigan. A review of machine learning applications in wildfire science and management. *Environmental Reviews*, 28(4):478–505, 2020.

Binh Thai Pham, Abolfazl Jaafari, Mohammadtaghi Avand, Nadhir Al-Ansari, Tran Dinh Du, Hoang Phan Hai Yen, Tran Van Phong, Duy Huu Nguyen, Hiep Van Le, Davood Mafi-Gholami, et al. Performance evaluation of machine learning methods for forest fire modeling and prediction. *Symmetry*, 12(6):1022, 2020.

Ioannis Prapas, Spyros Kondylatos, Ioannis Papoutsis, Gustau Camps-Valls, Michele Ronco, Miguel-Ángel Fernández-Torres, Maria Piles Guillem, and Nuno Carvalhais. Deep learning methods for daily wildfire danger forecasting. *arXiv preprint arXiv:2111.02736*, 2021.

Younes Oulad Sayad, Hajar Mousannif, and Hassan Al Moatassime. Predictive modeling of wildfires: A new dataset and machine learning approach. *Fire Safety Journal*, 104:130–146, 2019.

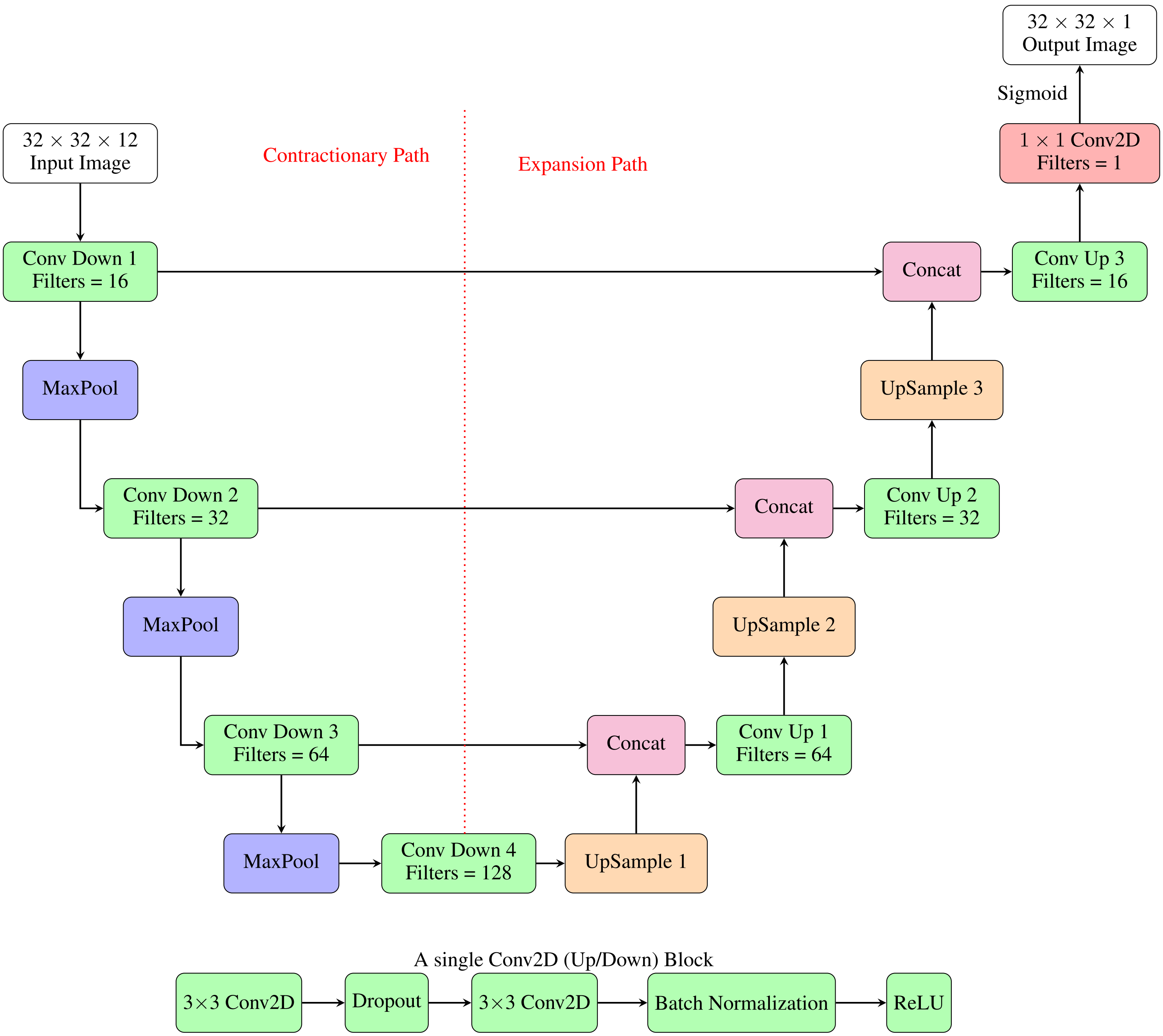

Figure 2: Deep learning model architecture: The top figure shows the U-Net model architecture for image segmentation. The bottom figure shows the layers of a single convolutional up/down block