THESIS

EXPLORING REMOTE SENSING DATA WITH HIGH TEMPORAL RESOLUTIONS FOR WILDFIRE SPREAD PREDICTION

Submitted by

Jack Fitzgerald

Department of Computer Science

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2024

Master's Committee:

Advisor: Nathaniel Blanchard

Nikhil Krishnaswamy
Dan Zimmerle

ABSTRACT

EXPLORING REMOTE SENSING DATA WITH HIGH TEMPORAL RESOLUTIONS FOR WILDFIRE SPREAD PREDICTION

The severity of wildfires has been steadily increasing in the United States over the past few decades, burning up many millions of acres and costing billions of dollars in suppression efforts each year. However, in the same few decades there have been great strides made to advance our technological capabilities. Machine learning is one such technology that has seen spectacular improvements in many areas such as computer vision and natural language processing, and is now being used extensively to model spatiotemporal phenomena such as wildfires via deep learning. Leveraging deep learning to model how wildfires spread can help facilitate evacuation efforts and assist wildland firefighters by highlighting key areas where containment and suppression efforts should be focused. Many recent works have examined the feasibility of using deep learning models to predict when and where wildfires will spread to, which has been enabled in part due to the wealth of geospatial information that is now publicly available and easily accessible on platforms such as Google Earth Engine. In this work, the First Week Wildfire Spread dataset is introduced, which seeks to address some of the limitations with previously released datasets by having an increased focus on geospatial data with high temporal resolutions. The new dataset contains weather, fuel, topography, and fire location data for the first 7 days of 56 megafires that occurred in the Contiguous United States from 2020 to 2024. Fire location data is collected by the Advanced Baseline Imager aboard the GOES-16 satellite, which provides updates every 5 minutes. Baseline experiments are performed using U-Net and ConvLSTM models to demonstrate some of the various ways that the First Week Wildfire Spread dataset can be used and to highlight its versatility.

ACKNOWLEDGEMENTS

I want to thank Dr. Nathaniel Blanchard for being a wonderful mentor and advisor. He has supported and guided me throughout my graduate studies, and I am eternally grateful for that. Joining his vision lab has been one of the best experiences, and I feel truly privileged to have met and worked with so many incredible people. I would not be where I am today without their support and the support from professors such as Nikhil Krishnasway, Craig Partridge, and Sangmi Pallickara. I want to say a special thank you to my peers Ethan Seefried and James Yost for their help in this work, and for being great friends. I also want to thank Daniel Zimmerle for being on my committee. Last but not least, I want to thank my friends and family for always being there for me and believing in me.

DEDICATION

I would like to dedicate this thesis to my parents, Kathleen and Patrick, who have always supported me and shown me so much love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

Wildfires in the United States have been increasing in severity over the past few decades, with an alarming trend that several millions of acres of land being burned each year is becoming the norm. Thousands of fires occur each year, with several hundred being classified as large fires. As millions of acres of land are burned each year, billions of dollars are being spent in the United States on suppression efforts [1]. As wildfires continue to increase in severity, wildland firefighters work diligently to provide an invaluable service to the public and making their jobs easier to perform should be a high priority. With the wealth of remote sensing data that are now provided via satellites, the prospect of understanding and predicting any wildfire's behavior is on the verge of becoming a reality. Predicting where wildfires will spread to can help firefighters plan better and allocate resources to areas that are most in need. Advanced techniques have and are continuing to be developed for predicting where any given wildfire will spread to, and one such exciting avenue is the potential for machine learning to perform this task. Machine learning has achieved impressive results in areas such as computer vision and natural language processing, and has shown promising results when modeling spatiotemporal phenomena. Exploring the application of machine learning for modeling wildfire spreading can lead to exciting new discoveries in model development, and more importantly can help humanity in dealing with natural disasters that are becoming evermore prevalent.

Wildfires are complex phenomena that can be difficult to model due to the huge variety of factors that influence their behavior. The spatial distribution of fuels, changes in weather, and topography are all unique characteristics of the environments that wildfires can occur in. Many of the early, and even current, approaches to modeling the spreading of wildfires are based on complex mathematical models, such as the Rothermel model [2]. The Rothermel model can be used to calculate the rate of spread of a fire, and is used as the basis in prediction models such as FARSITE [3]. Besides physically modeling fires, many different approaches have been developed to predict wildfire spreading including using AI, machine learning, and networks of IoT sensors. Cellular automata has been used to model wildfire spreading behavior [4–7], which model wildfires as a system of cells that interact with each other and simultaneously evolve over time. Genetic algorithms have been used to optimize physical simulations of wildfires [8–10] and to develop suppression strategies [11]. Networks of wireless sensor devices have been used to predict wildfire spreading [12,13] and to monitor the spread of wildfires [14]. Using wireless sensors on the ground provides a useful alternative to satellites by having a much higher sampling rate.

Machine learning has been used extensively in wildfire related tasks for the past two decades [15], including being used to predict wildfire spreading. Early approaches used several machine learning approaches such as state vector machines [16], artificial neural networks [17], logistic regression [18], and random forest [19]. However, deep learning has recently gained a significant amount of traction in being applied to predicting wildfire spreading, likely because remote sensing data related to wildfire spreading is now easily accessible and widely available thanks to platforms such as Google Earth Engine (GEE). Several new benchmarks and datasets have been created for facilitating the development of wildfire spread prediction models [20–23], and there have been many deep learning models developed which utilize convolutional [24–26] and transformer [27] architectures.

1.1 Wildfire Spread Prediction Datasets

Although many deep learning models have been created for wildfire spread prediction, several works do not publicly release the dataset that they used to train and test their models [28–30]. However, to alleviate this problem, there have been several publicly released datasets for training deep learning models to predict wildfire behavior. The most notable public dataset is the NEXTDAYWILDFIRESPREAD [20]. This dataset uses GEE to provide samples in the form of 64 x 64 images that contain geospatial information, including a previous fire mask that can be fed to a model that is then asked to predict the next fire mask that is 24 hours ahead. The NEXTDAYWILDFIRESPREAD dataset has inspired many works since its release [31–33], and has led to the creation of new public datasets such as WILDFIRESPREADTS [21] that seeks to address some of its limitations.

1.1.1 Next Day Wildfire Spread

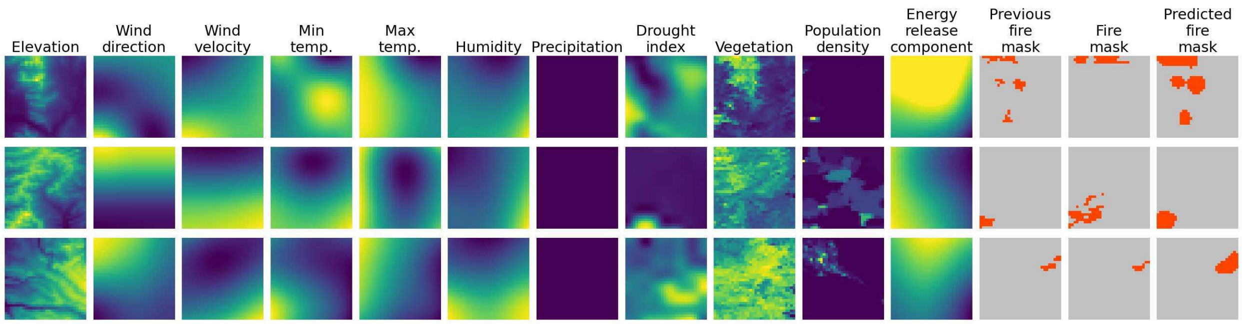

Figure 1.1: Some samples from the NEXTDAYWILDFIRESPREAD dataset with predictions made by the Wildfire Prediction Network (WPN) model from [32]

The NEXTDAYWILDFIRESPREAD dataset from [20] utilized the MOD14A1.061 fire products [34] from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensors aboard the Terra satellite. Weather data is collected from the GRIDMET dataset [35] in the form of minimum and maximum temperature, wind direction and speed, specific humidity, energy release component, precipitation, and drought indices. Elevation data comes from the Shuttle Radar Topography Mission (SRTM) [36]. Vegetation indices (NDVI) are provided by the VNP13A1 dataset [37] collected by VIIRS sensors aboard the Suomi National Polar-orbiting Partnership (S-NPP). Finally, population data is provided from the Gridded Population of the World v4 (GPWv4) dataset [38]. Figure 1.1 shows some of the samples from the NEXTDAYWILDFIRESPREAD dataset along with predictions made by the Wildfire Prediction Network from [32].

The NEXTDAYWILDFIRESPREAD dataset contains 18,545 samples. Each sample can be treated as an input image with 11 channels, where each channel represents one of the bands in Table 1.1, and an output image with 1 channel that represents the fire mask of the fire 24 hours later. The size of each sample is 64 x 64 pixels and represents a 64km x 64km geospatial region since each pixel represents a 1km x 1km region, which is due to the spatial resolution of the Terra MODIS

Table 1.1: A summary of each of the datasets comprising the NEXTDAYWILDFIRESPREAD dataset. the SRTM dataset only provides a single image from around February of 2000.

Datasets	Temporal Res.	Spatial Res.	Bands
Terra MOD14A1.061	daily	1km	Fire mask
GRIDMET	daily	4.5km	Minimum temperature (tmmn)
			Maximum temperature (tmmx)
			Specific humidity (sph)
			Wind direction (th)
			Wind velocity at 10m (vs)
			Energy release component (erc)
			Precipitation amount (pr)
SRTM	Feb. 2000	90m	Elevation
VNP13A1	8-days	500m	Normalized Difference Vegetation Index (NDVI)
GPWv4	5 years	1km	Population density

fire products being 1km. However, the main experiments performed by the authors of the NEXTDAYWILDFIRESPREAD dataset were done on randomly cropped 32 x 32 pixel sections because the original 64 x 64 pixel samples had fires that were located almost entirely in the center of the frames.

1.1.2 WildfireSpreadTS

The NEXTDAYWILDFIRESPREAD dataset only contains samples that have a single previous fire mask, which means that the model has no knowledge of how the fire has evolved before it is asked to make a prediction. To address this lack of temporal data in the NEXTDAYWILDFIRESPREAD dataset, the WILDFIRESPREADTS dataset was created with the inclusion of samples that include multiple fire masks from previous days. Furthermore, it includes a total of 23 features by including more topographic features, more fuel features, and weather forecasts. Table 1.2 provides descriptions of the features that are included in the WILDFIRESPREADTS dataset. WILDFIRESPREADTS provides 13,607 images gathered from 607 wildfires in the United States, which were selected from the GlobFire dataset [39], that occurred during January 2018 to October 2021. The dataset also provides full coverage of the 607 wildfires that it includes, meaning it contains images from each day of the fires. The images themselves ranged from 304 x 207 pixels to 356 x 308 pixels.

Table 1.2: A summary of each of the datasets comprising the WILDFIRESPREADTS dataset. It utilizes the VIIRS active fire products, which have greater spatial and temporal resolutions compared to the MOD14A1.061 fire products used in the NEXTDAYWILDFIRESPREAD dataset. * The 2-4 daily images that are provided by the VIIRS active fire products get combined into a single image for each day. ** The aspect and slope features are calculated based on the elevation data.

Datasets	Temporal Res.	Spatial Res.	Bands
VIIRS Active Fire	2-4 per day*	375m	Fire mask
VNP09GA	daily	500m	I1
			I2
		1000m	M11
VNP13A1	8-days	500m	Normalized Difference Vegetation Index (NDVI)
			2 band Enhanced Vegetation Index (EVI2)
MCD12Q1.061	Yearly	500m	IGBP land cover type (LC_Type1)
GRIDMET	daily	4.5km	Minimum temperature (tmmn)
			Maximum temperature (tmmx)
			Specific humidity (sph)
			Wind direction (th)
			Wind velocity at 10m (vs)
			Energy release component (erc)
			Precipitation amount (pr)
GFS	hourly	27.83km	Wind speed
			Wind direction
			Mean temperature
			Total precipitation
			Specific humidity
SRTM	Feb. 2000	90m	Elevation
			Aspect**
			Slope**

WILDFIRESPREADTS utilizes the VIIRS active fire products [40] to provide the fire masks at a spatial resolution of 375m, and aggregates the 2-4 images per day into a single image. Bands I1, I3, and M11 from the VIIRS surface reflectance product (VNP09GA) [41] are used for cloud, smoke, and vegetation detection. The VIIRS VNP13A1 dataset [37] provides the NDVI and EVI2 vegetation indices. Weather data is provided by GRIDMET [35], and weather forecasts are provided by the Global Forecast System (GFS) [42]. Land cover data is provided yearly from the MODIS MCD12Q1.061 dataset [43] at a spatial resolution of 500m. Finally, elevation data is provided by the NASA SRTM dataset [36], and the aspect and slope are calculated based on the elevation data.

1.1.3 Mesogeos

MESOGEOS [22] is a multipurpose dataset designed for perform multiple different tasks related to modeling wildfires. Currently, MESOGEOS has two tracks for training and testing models on fire danger and burned area forecasting, but it can be extended with other tracks such as fire size prediction and wildfire susceptibility mapping. Researchers are encouraged to use MESOGEOS to train their models and submit their results to the tracks. There are 27 features included in the MESOGEOS dataset, which includes topographic, vegetation, weather, and anthropogenic data. Similar to NEXTDAYWILDFIRESPREAD, the MODIS active fire products [34] are used, except it's only used to estimate ignition cells and ignition dates. The burned area data comes from the European Forest Fire Information System (EFFIS).

1.1.4 EO4WildFires

The EO4WILDFIRES dataset [23] combines meteorological data, images from Sentinel-1 and Sentinel-2, and fire masks from EFFIS for 31,742 fire events. The authors perform experiments with ResNet-32 to predict the total size of the burned area from a wildfire based on its starting ignition point. Further experiments are performed on the EO4WILDFIRES dataset in [44], where various segmentation models are used to attempt to generate masks of the total burned area from the wildfires.

1.2 Creating a New Dataset

While there are already several publicly available datasets for modeling wildfire behavior, none appear to make use of the fire/hot spot characterization products from the Advanced Baseline Imager (ABI) aboard the GOES-R series of satellites (GOES-16, GOES-17, GOES-18, and GOES-19) [45, 46]. The ABI aboard GOES-16 is able to provide data at an astounding temporal resolution of 5 minutes for areas that are within the Contiguous United States, albeit at a relatively low spatial resolution of 2km. The high temporal resolution of the ABI can be used to fill in the gaps within the day that sensors such as MODIS and VIIRS are unable to collect. Utilizing the fire products from the GOES-16 ABI, we can develop a new dataset that is very similar to the previously described datasets such as NEXTDAYWILDFIRESPREAD and WILDFIRESPREADTS, but with a much higher temporal resolution. In this work, a new dataset called FIRSTWEEKWILDFIRESPREAD is introduced, which uses the fire products from the ABI aboard the GOES-16 satellite. The FIRSTWEEKWILDFIRESPREAD dataset can be used to perform similar experiments that can be done with the NEXTDAYWILDFIRESPREAD and WILDFIRESPREADTS datasets, but it can also be used to train machine learning models that make wildfire spread predictions at many different time steps. FIRSTWEEKWILDFIRESPREAD is described in detail in Chapter 2, and how the dataset was collected is explained in Chapter 3.

1.3 Contributions

The contributions of this work can be summarized as follows:

- A new dataset called FIRSTWEEKWILDFIRESPREAD is introduced that utilizes the fire products from the ABI sensor aboard the GOES-16 satellite, and is structured so that a wide range of experiments can be performed. This new dataset provides fire masks with a temporal resolution of 5 minutes, which is substantially higher than the temporal resolution of the fire masks provided by the satellites containing MODIS and VIIRS sensors.

- Baseline experiments are performed on FIRSTWEEKWILDFIRESPREAD to demonstrate the variety of experiments that can be performed on it. It can be used to perform similar experiments to the NEXTDAYWILDFIRESPREAD dataset, where models are trained to predict 24 hours ahead, but at significantly more time offsets. Moreover, FIRSTWEEKWILDFIRESPREAD can be used to construct time series samples to train and test models that are designed to handle spatiotemporally evolving phenomena such as wildfires.

Chapter 2

The First Week Wildfire Spread Dataset

Inspired by the work performed in [20] with the NEXTDAYWILDFIRESPREAD dataset, the FIRSTWEEKWILDFIRESPREAD dataset is created with an increased focus on providing data with high temporal resolutions. The FIRSTWEEKWILDFIRESPREAD dataset is structured similarly to NEXTDAYWILDFIRESPREAD, but it also differs in several key ways. The primary difference is the inclusion of data with higher temporal resolutions, but some of the other differences seek to address limitations of NEXTDAYWILDFIRESPREAD. Some of the limitations in NEXTDAYWILDFIRESPREAD include a lack of metadata describing the wildfires, it is unknown which coordinate reference system was used, the images are of relatively small spatial regions, and it has an inflexible design that limits the variety of experiments that can be performed.

The Terra MODIS fire products used by the NEXTDAYWILDFIRESPREAD dataset are only updated every 24 hours, which can be a potential downside when considering that wildfires can evolve rapidly over the course of a day. The GOES-16 satellite can provide updates on fire locations every five minutes, albeit at a reduced spatial resolution of 2km. Creating a dataset similar to the NEXTDAYWILDFIRESPREAD dataset, but using the GOES-16 satellite, can capture nearly the entire evolution of a wildfire with much fewer gaps. A wildfire spread prediction dataset that utilizes the GOES-16 satellite could further help facilitate the development and analysis of various machine learning models that can be applied to wildfire spread prediction.

There are currently several satellites equipped with sensors that are capable of detecting fires. This includes the Terra and Aqua satellites equipped with the Moderate Resolution Imaging Spectroradiometer (MODIS) [47], the NOAA-20, NOAA-21, and Suomi-NPP satellites equipped with the Visible Infrared Imaging Radiometer Suite (VIIRS) [48], and the GOES-R series of satellites equipped with the Advanced Baseline Imager (ABI) [45, 46]. NOAA-20, NOAA-21, and Suomi-NPP form a Joint Polar Satellite System (JPSS) and provide full coverage of the earth twice a day, NOAA-20 follows behind Suomi-NPP by about 50 minutes. The GOES-R series of satellites

(GOES-16, GOES-17, GOES-18, and now GOES-19) are able to provide a high temporal resolution because they are geostationary. The GOES-16 satellite in particular is able to provide fire detection products every 5 minutes for the entire Contiguous United States (CONUS). Leveraging the high temporal resolution of the GOES-16 satellite can help us better understand how fires evolve. The one downside to the GOES satellites is that they have a relatively low spatial resolution of 2km, meaning that each pixel represents a 2km x 2km region. In contrast, MODIS sensors provide a 1km spatial resolution and the VIIRS sensors provide a 375m resolution. However, having a high temporal resolution means that we can know more about where the fire has spread to, which can be crucial information due to how quickly fires can spread. Having only one or two updates a day on where a fire is burning can theoretically result in significant knowledge gaps of how the fire has evolved, potentially impacting the predictions of where the fire is most likely to spread to.

The FIRSTWEEKWILDFIRESPREAD dataset utilizes GOES-16 FDCC to provide fire detection products, GOES-16 MCMIPC to provide RGB images, Real-Time Mesoscale Analysis (RTMA) to provide weather data, NASA Digital Elevation Model (NASADEM) for elevation data, and the Terra and Aqua combined MODIS Land Cover Type (MCD12Q1) Version 6.1 to provide land cover types. GOES-16 MCMIPC only provides dedicated red and blue bands, but it does not have a dedicated green band. The green band is a combination of the red (CMI_C02), blue (CMI_C01), and the near-IR veggie (CMI_C03) bands, and is calculated with the following equation from [49],

$$\text{CMI\_GREEN} = 0.45 * \text{CMI\_C02} + 0.45 * \text{CMI\_C01} + 0.1 * \text{CMI\_C03}. \tag{2.1}$$

This gives us something close to a true-color RGB image. A short summary of each dataset used in FIRSTWEEKWILDFIRESPREAD is shown in Table 2.1. In contrast, NEXTDAYWILDFIRESPREAD uses MODIS sensors to detect fire anomalies and the GRIDMET dataset for weather, both of which have daily temporal resolutions. Having more up to date weather information with the RTMA weather dataset, which has an hourly temporal resolution, can be important in understanding the behavior of a fire. The FIRSTWEEKWILDFIRESPREAD dataset does not contain information about population, vegetation, or energy release component, but instead includes land cover types from

MCD12Q1.061 and RGB images from GOES-16 MCMIPC. Figure 2.1 shows 3 samples from FIRSTWEEKWILDFIRESPREAD to illustrate what the features look like.

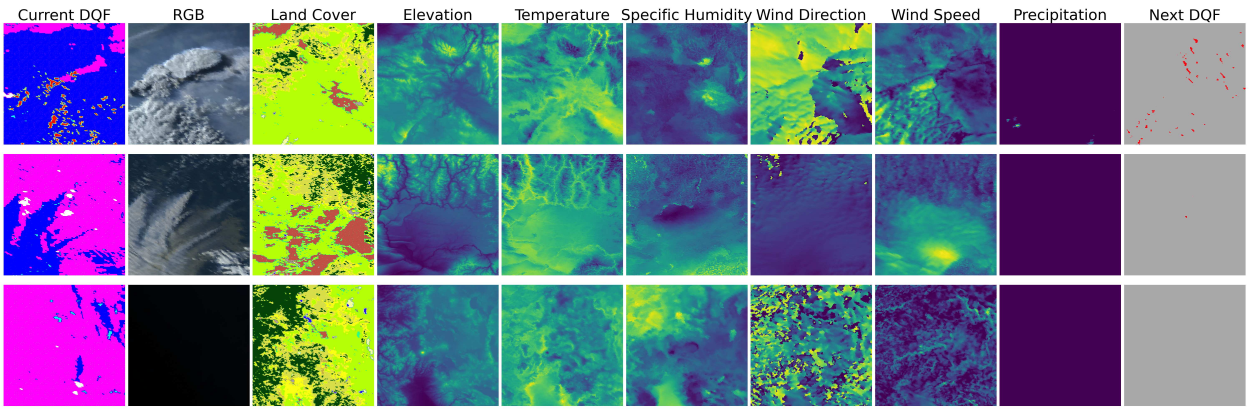

Figure 2.1: Some samples from the FIRSTWEEKWILDFIRESPREAD dataset. The Next DQF fire masks are 24 hours ahead of the current DQF fire masks.

Table 2.1: A summary of each of the datasets comprising the ***First Week Wildfire Spread*** dataset. the NASADEM dataset only provides a single image from around February of 2000. * The green band from GOES-16 MCMIPC is a combination of the red, blue, and near-IR veggie bands.

Datasets	Temporal Res.	Spatial Res.	Bands
GOES-16 FDCC	5 minutes	2km	DQF
GOES-16 MCMIPC			RGB*
RTMA	1 hour	2.5km	Temperature (TMP)
			Specific humidity (SPFH)
			Wind direction (WDIR)
			Wind speed (WIND)
			Precipitation (ACPC01)
NASADEM	Feb. 2000	30m	Elevation
MCD12Q1.061	Yearly	500m	IGBP land cover type (LC_Type1)

The FIRSTWEEKWILDFIRESPREAD dataset captures data from the first week of fires, potentially before many suppression efforts have occurred which was a concern pointed out in [20]. In total, the FIRSTWEEKWILDFIRESPREAD dataset contains data from 56 of the largest wildfires across CONUS.

One of the main improvements to the FIRSTWEEKWILDFIRESPREAD dataset is the inclusion of metadata for each wildfire. A downside to NEXTDAYWILDFIRESPREAD is that it is unclear exactly where each sample comes from due to the lack of any metadata, thus making replication efforts more difficult. Information on where the wildfires were, when they started, their perimeters, and how large they were was obtained from the Wildland Fire Interagency Geospatial Services (WFIGS) Interagency Fire Perimeters (IFP) dataset. This dataset is in JSON format and contains geometry data describing the most up to date perimeter for each fire within the dataset, which is usually the final perimeter of the fire unless it is currently ongoing. The IFP dataset also provides a plethora of metadata for each wildfire, including the date it was discovered, the name of the fire, its initial starting location, etc. All of the metadata, including the perimeters, is included with each of the wildfires in the FIRSTWEEKWILDFIRESPREAD dataset. The total area of the wildfires was calculated by finding the total areas of the perimeters, as opposed to using the value of the final size in acres reported in the metadata for each fire. This is because it's unclear how the final size in acres that was reported in the metadata was calculated.

Chapter 3

Methodology

All data for the FIRSTWEEKWILDFIRESPREAD dataset was collected from Google Earth Engine (GEE), using the Python package called `earthengine-api`. Experiments are performed on the newly collected FIRSTWEEKWILDFIRESPREAD dataset and compared to some of the experiments performed on the NEXTDAYWILDFIRESPREAD dataset in [32]. A model based on the U-Net architecture [50,51], which is known for its ability to perform well in semantic segmentation tasks, is used to predict which regions a wildfire will spread to based only on current knowledge of where the fire is along with the most up to date weather, topography, and land cover information. Furthermore, a convolutional LSTM (ConvLSTM) model [52] is used to process a time series of wildfire information to make wildfire spread predictions.

3.1 Google Earth Engine Data Collection

Similar to NEXTDAYWILDFIRESPREAD, all data for the FIRSTWEEKWILDFIRESPREAD dataset was collected from GEE datasets using the `earthengine-api` Python package. With the `earthengine-api`, users can do various transformations on a plethora of satellite imagery, and then download it to their local machines to perform further analysis tasks. Several datasets from GEE are utilized to create the FIRSTWEEKWILDFIRESPREAD dataset including GOES-16 FDCC, GOES-16 MCMIPC, RTMA, NASADEM, and MCD12Q1.061, which all provide spatiotemporal information about regions of interest within the Contiguous United States (CONUS).

In order to collect data about wildfires using `earthengine-api`, we first find the coordinates and perimeters of wildfires that match a certain criteria. Then we determine the regions of the satellite images based a wildfire's coordinates, the wildfire's start and end dates, and the coordinate reference system (CRS) that we want to project all of the satellite images to. The `earthengine-api` package handles all of the region calculations and CRS projections under the hood so that users don't have to worry about it. Finally, `earthengine-api` will generate an `ImageCollection` object which can be used to download all of the wildfire images for a specific region that are within the start and end dates.

The wildfire perimeters were collected from the Wildland Fire Interagency Geospatial Services (WFIGS) Interagency Fire Perimeters (IFP) dataset from the National Interagency Fire Center (NIFC), which contains the perimeters of wildfires and their associated metadata for more than 26,000 wildfires that have occurred in the United States. The selection criteria for a wildfire was that it must have a total burned area greater than 100,000 acres, it occurred after May 24th, 2017, and that it must not have occurred in the states of Alaska or Hawaii. The earliest date that GEE provides GOES-16 data is May 24th, 2017, and the sensors onboard the GOES-16 satellite are unable to capture Alaska and Hawaii at a temporal resolution of 5 minutes (the GOES-16 Full-Disk dataset does provide data for these states, albeit at a temporal resolution of 10 minutes). There were 56 wildfires that matched the specified criteria, and they all occurred during or after 2020. However, the WFIGS IFP dataset is not comprehensive and might not contain data for older fires that match this criteria.

After selecting wildfires from the WFIGS IFP dataset, images from each of the GEE datasets could then be collected. The images were collected by first calculating the centroid (the central latitude and longitude coordinates) of the perimeter geometry for each selected fire, then creating a bounding box with an approximate size of 400km x 400km around the center point of each fire. The images from each of the GEE datasets were scaled so that each pixel in the images represents a 2km x 2km geospatial region, i.e. the spatial resolution is 2km, which resulted in the images having a size of approximately 200 x 200 pixels. However, the images are not exactly 200 x 200 pixels due to the fact that the kilometers in a degree of longitude is a function of latitude. A degree of longitude has the most kilometers at the equator and decreases as the latitude moves towards the poles. The NDWS dataset contains images of 64km x 64km geospatial regions, although at a much higher spatial resolution of 1km, and thus a major difference in FIRSTWEEKWILDFIRESPREAD is that the area taken into consideration for each image is more than four times that of the NEXTDAYWILDFIRESPREAD images.

All of the images were projected to the EPSG:4326 CRS, which is also known as the World Geodetic System 1984 (WGS84). Different types of map projections can produce different looking images, which could potentially have an impact on model inferences [53] if the images that are passed to the model are not in the projection of the images that the model was trained on. However, if the CRS of a particular geospatial image is known, then it can be transformed into the CRS that the model expects and has been trained on [54]. One of the limitations of the NEXTDAYWILDFIRESPREAD dataset is that the CRS used is not specified, and thus it would not be possible to project new data to the CRS that the model was trained on.

Finally, all of the images were downloaded locally by using the `Image.getDownloadURL` function in the `earthengine-api` package which takes a region, start and end dates, scale, and CRS, and then returns a URL which can be used to download a GeoTIFF file. The start date was collected from the `attr_FireDiscoveryDateTime` field in the WFIGS IFP metadata, and the end data was computed to be one week after the start date. The returned URL was then used to download the GeoTIFF file over HTTPS via the Python `requests` package.

3.2 Dataset Structure

The structure of the FIRSTWEEKWILDFIRESPREAD dataset is designed so that a variety of experiments can be performed. In FIRSTWEEKWILDFIRESPREAD, each fire has its own directory that is named after its Integrated Reporting of Wildfire Information (IRWIN) ID. Within each fire's directory, there is a metadata file and several directories for storing the GeoTIFF files downloaded from the GEE datasets. The metadata file is named `fire-metadata.json` and it contains all of the fire's information included in the WFIGS IFP dataset, including its perimeter data. Within a fire's directory there are also individual directories for each GEE dataset (GOES-16 FDCC, GOES-16 MCMIPC, RTMA, NASADEM, and MCD12Q1.061) that contain the downloaded GeoTIFF files.

The filenames of the GeoTIFF files include the dates in UTC time, so that they can be used to create time series samples and adjust how far into the future a prediction should be. For example, for each GeoTIFF image in the GOES-16 FDCC directories, we can attempt to find a corresponding GeoTIFF image that is an hour ahead, and then we can use the GOES-16 FDCC image that is an hour ahead as the ground truth label that a model can attempt to predict. In practice, any time offset into the future that is a multiple of 5 minutes in the range of 5 - 10,080 (10,080 is the total amount of minutes in a week) can be used as the ground truth label for a particular GOES-16 FDCC image.

3.3 Experiment Process

The problem of predicting wildfire spreading based on remote sensing imagery is very similar to semantic segmentation tasks that are commonly seen in computer vision [55]. Each band collected from the GEE datasets can be represented as a channel in an image, similar to how red, green, and blue channels comprise images that are visible to the human eye. This makes it easy to apply convolutional neural networks (CNNs) to the data. Following in the footsteps of NEXTDAYWILDFIRESPREAD, in which the authors performed experiments on their data using semantic segmentation models [20], the FIRSTWEEKWILDFIRESPREAD dataset takes a similar approach and is intended to be used for training models that are capable of performing semantic segmentation. Several unique baseline experiments are performed to highlight the capabilities of FIRSTWEEKWILDFIRESPREAD, including an experiment that measures the performance of a U-Net model developed in [32], dubbed the Wildfire Prediction Network (WPN), as it attempts to make predictions at increasingly higher time offsets starting from 5 minutes ahead up 24 hours ahead. A second similar experiment is performed with a convolutional long short term memory (ConvLSTM) model [52] to demonstrate the potential of the FIRSTWEEKWILDFIRESPREAD dataset to be used for time series analyses.

3.3.1 Data Pre-Processing

Before the WPN and ConvLSTM models could be trained and tested, several pre-processing steps needed to be made. First, the wildfires needed to be separated into training, testing, and validation splits. Then some of the features from the data were adjusted by converting to different units, clipping outliers to specific values, converting 16-bit values to 8-bit values, etc. Finally, the mean and standard deviations were calculated for each of the features based on the training split so that they can be standardized.

The training, testing, and validation splits consisted of unique fires to ensure that there was no overlap between them to avoid any similarities between the splits that could undermine the testing results. In the NEXTDAYWILDFIRESPREAD dataset, it's unclear which samples came from which fire, and so there is a potential that samples from the same fires are included in all of the splits. When making 24 hour predictions, including samples from the same fire might not have a significant influence on the results due to the stochastic nature of wildfires from day-to-day. Nonetheless, it's likely safer to avoid any potential overlaps between the splits that could undermine the integrity of the model results. For the FIRSTWEEKWILDFIRESPREAD dataset, there was a roughly 70-15-15 split for the training, testing, and validation sets respectively which meant that the training set had 40 unique fires, and the testing and validation sets had 8 unique fires each, all of which were randomly binned into each set.

The features in FIRSTWEEKWILDFIRESPREAD that were adjusted were temperature, elevation, the RGB bands from GOES-16 MCMIPC, and the ground truth fire mask from GOES-16 FDCC. The temperature values were converted from Celsius to kelvin by simply adding 273.15 to each temperature value. Any negative elevation values were clipped to 0, however there are places below sea level that are not underwater such as Death Valley National Park. The red, green, and blue channels from GOES-16 MCMIPC are converted from 16-bit values to 8-bit values by dividing the maximum value for each channel in the training split so that each value is in the range [0, 1], then each value is multiplied by 255 to ensure that all RGB values are in the 8-bit range of [0, 255]. Since the maximum value from the training set is used to convert RGB values in the testing and validation sets from 16-bit to 8-bit, which may have maximum RGB values greater than those seen in the training set, any values that are greater than 255 are clipped to 255. The ground truth fire masks are converted by changing all of the fire pixels (which are normally 0) to 1, and all of the other pixels to 0, this makes it easier to perform binary classification on the pixels during training and testing. In the ground truth fire masks and the model predictions, a value of 1 represents the presence of fire and 0 represents no fire.

All data is standardized before being passed into the U-Net and ConvLSTM models. This is done by first calculating the mean and standard deviation of all features in the training split, and then subtracting each features value by its corresponding mean and dividing by its corresponding standard deviation. This is known as the z-score standardization process and can be described by

$$\mathbf{z} = \frac{x - \bar{\mathbf{x}}}{\mathbf{s}}, \tag{3.1}$$

where $\mathbf{z}$ is the standardized feature value, x is the original feature value, $\bar{\mathbf{x}}$ is the sample mean of the feature in the training split, and $\mathbf{s}$ is the sample standard deviation of the feature in the training split. This process ensures that each feature in the training split has a mean of 0 and a standard deviation of 1, making it so that the values across all features are on a similar scale. This standardization process is also applied to all the feature values in the testing and validation splits, which likely have different sample means and standard deviations, but the sample means and standard deviations from the training split are hopefully good estimations of the population means and standard deviations.

Calculating the means and standard deviations is a straight forward process for all of the features except for wind direction, this is because wind direction is a directional statistic [56, 57] and requires a different method for calculating the mean and standard deviation. The values of the wind direction are in degrees, and naively attempting to calculate the mean by summing all values and dividing by the total elements can result in a mean value that does not reflect what the true direction of the mean value should be. For example, if we have a sample with two northward pointing wind direction values of 1° and 359°, then intuitively the mean value should be 0°. But with the naive mean calculation the mean will be 180° which is pointing south. Therefore, we need a formula for calculating the mean wind direction that will preserve the directions of the values in the sample.

This can be done by first converting the wind direction values into 2D vectors by using

$$r_i = \begin{pmatrix} \cos \alpha_i \\ \\ \sin \alpha_i \end{pmatrix}, \tag{3.2}$$

where α_i is the value in radians for wind direction i, and r_i is the unit vector for wind direction i. The mean wind direction for the training split can then be calculated by

$$\bar{r} = \frac{1}{N} \sum_i r_i, \tag{3.3}$$

where $\bar{r}$ is the sample mean resultant vector, and N is the total number of elements in the sample. We can then find the angle of $\bar{r}$ with

$$\bar{r}_{angle} = \begin{cases} \arctan \frac{\bar{r}_{sine}}{\bar{r}_{cosine}} & \bar{r}_{cosine} \geq 0 \\ \\ \arctan \frac{\bar{r}_{sine}}{\bar{r}_{cosine}} + \pi & \bar{r}_{cosine} < 0 \end{cases} \tag{3.4}$$

We can then use equation (3.2) to convert all of the wind direction pixel values from the training split to unit vectors, and then use equation (3.3) and equation (3.4) to find the angle of the mean resultant vector from those unit vectors.

Calculating the standard deviation for a directional statistic typically involves using the length of the mean resultant vector $\bar{r}$, but this is not sufficient for standardizing the wind direction using equation (3.1) because it is not an angular measure. We need an equation for calculating the standard deviation of the wind direction that will give us a value with units in radians, a rudimentary way of performing such a calculation can be given by

$$\mathbf{s_{dir}} = \frac{1}{N} \sum_i \arccos\left(\cos\left(\alpha_i - \bar{r}_{angle}\right)\right). \tag{3.5}$$

which will return a value in the range of $[0, \pi]$ degrees, thus making $\mathbf{s_{dir}}$ represent the average distance from the mean angle $\bar{r}_{angle}$. Then we can use the following equation,

$$\mathbf{z_{dir}} = \frac{(\alpha - \bar{r}_{angle} + \pi) \bmod 2\pi - \pi}{\mathbf{s_{dir}}}, \tag{3.6}$$

to standardize the wind direction values.

3.3.2 U-Net Wildfire Prediction Network

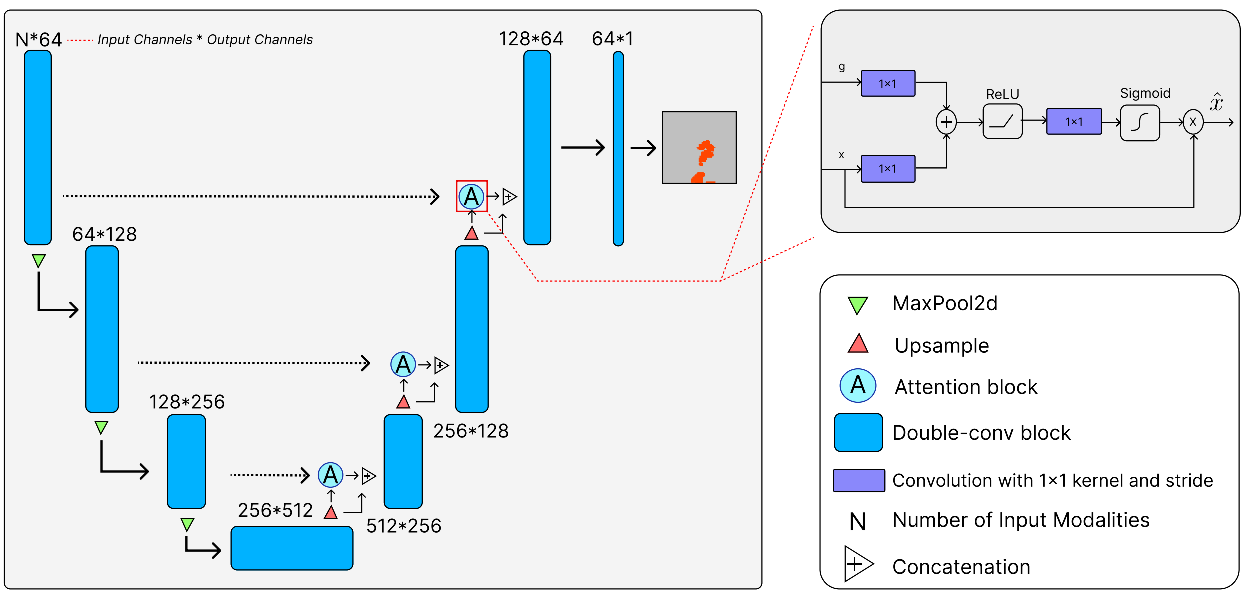

Figure 3.1: The architecture of the Wildfire Prediction Network (WPN) from [32], which is based on the Attention U-Net model [51]

The Wildfire Prediction Network (WPN) is used as a baseline model in the experiments on FIRSTWEEKWILDFIRESPREAD, and it was first developed in [32] where it was applied to NEXT-DAYWILDFIRESPREAD. The WPN model is based on the U-Net architecture, which was originally developed to perform semantic segmentation on biomedical images [50]. Several variants of the original U-Net architecture have been developed since its inception [51, 58–60]. In particular, the WPN model is an implementation of the Attention U-Net architecture from [51] and is described in Figure 3.1. Similar to the experiments performed in [32], where the WPN model was used on NEXTDAYWILDFIRESPREAD to make wildfire spread predictions that are 24 hours ahead of the current known fire locations, the WPN model is used on FIRSTWEEKWILDFIRESPREAD to make wild spread predictions at increasingly higher time offsets. In this work, the WPN model is trained to make spread predictions that are 5, 30, 60, 120, 180, 360, 720, and 1440 minutes ahead, both demonstrating the ability of FIRSTWEEKWILDFIRESPREAD to perform wildfire spread prediction experiments at various time offsets and analyzing how well the WPN can make predictions that are less than 24 hours ahead.

3.3.3 Convolutional LSTM

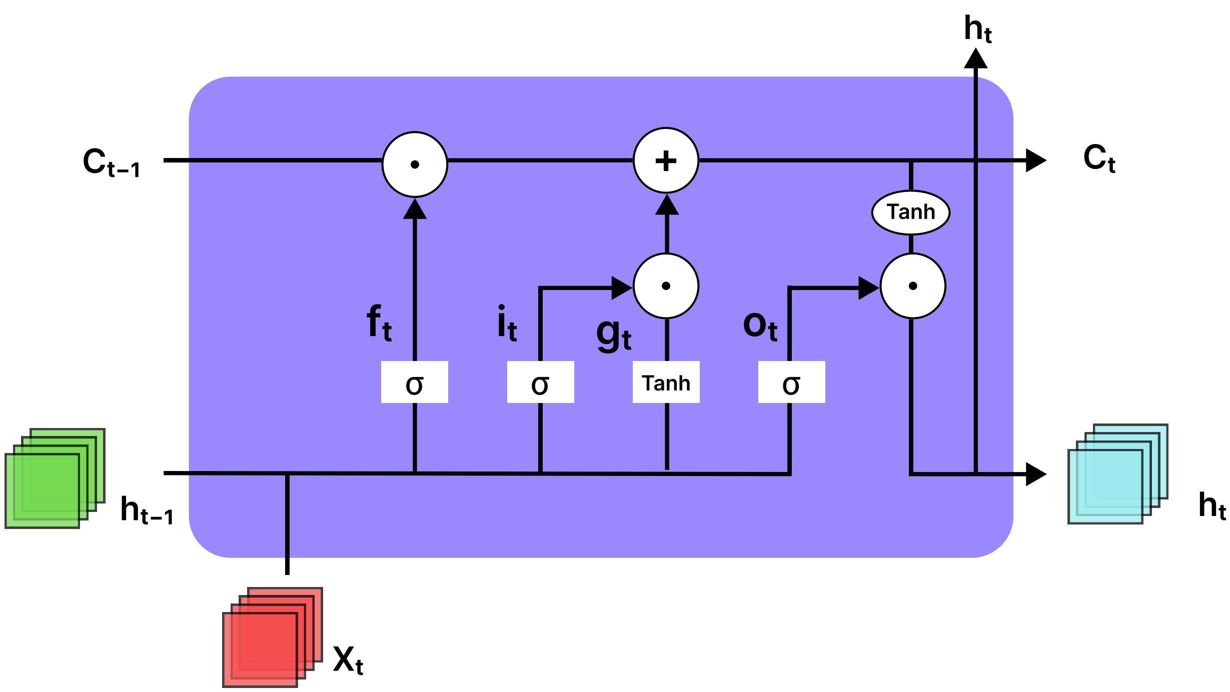

Figure 3.2: A depiction of a long short-term memory (LSTM) cell in the ConvLSTM model. The input X_t is concatenated with the hidden cell h_{t-1} from the previous time step, and is convolved by four separate gates, f_t, i_t, g_t, and o_t which represent the forget, input, control, and output gates respectively. The values in the hidden cell can be used as the output of the model, after they have passed through the LSTM cell. For example, h_t can be used as the output of the ConvLSTM model at time step t, which represents the prediction for time $t + 1$.

FIRSTWEEKWILDFIRESPREAD is also intended to be used for time series analyses. To demonstrate this potential, a convolutional LSTM (ConvLSTM) model [52] is used to perform wildfire spread predictions while taking into account several previous updates on the wildfire locations at time steps t, $t-1$, $t-2$, ..., $t-n$. The ConvLSTM is structurally very similar to a normal LSTM model [61], but instead of fully connected layers it uses convolutional layers on the concatenation of the inputs and hidden states. The ConvLSTM was originally developed for modeling spatiotemporal phenomena, and is thus a suitable baseline to use for predicting wildfire spreading. The baseline ConvLSTM used in the experiment contains only a single LSTM cell, which takes in 5 sequential inputs at $t-4$, $t-3$. $t-2$. $t-1$, and t, and tries to predict where the wildfire has spread to at $t+1$. This is similar to the experiment with WPN, but it explores the potential that passing in multiple previous wildfire locations has on model performance. In theory, if the model is aware of areas that the wildfire has already spread to, then it may learn more useful patterns, such as knowing which areas have already been burned, or if the trajectory of the fire is tending to spread in specific directions. The equations for the ConvLSTM model can be given by

$$\mathbf{f_t} = \sigma(Conv(\mathbf{x_t}; \mathbf{w_{xf}}) + Conv(\mathbf{h_{t-1}}; \mathbf{w_{hf}}) + \mathbf{b_f})$$

$$\mathbf{i_t} = \sigma(Conv(\mathbf{x_t}; \mathbf{w_{xi}}) + Conv(\mathbf{h_{t-1}}; \mathbf{w_{hi}}) + \mathbf{b_i})$$

$$\mathbf{o_t} = \sigma(Conv(\mathbf{x_t}; \mathbf{w_{xo}}) + Conv(\mathbf{h_{t-1}}; \mathbf{w_{ho}}) + \mathbf{b_o})$$

$$\mathbf{g_t} = \tanh(Conv(\mathbf{x_t}; \mathbf{w_{xg}}) + Conv(\mathbf{h_{t-1}}; \mathbf{w_{hg}}) + \mathbf{b_g})$$

$$\mathbf{c_t} = \mathbf{f_t} \odot \mathbf{c_{t-1}} + \mathbf{i_t} \odot \mathbf{g_t} \tag{3.7}$$

$$\mathbf{h_t} = \mathbf{o_t} \odot \tanh(\mathbf{c_t}) \tag{3.8}$$

where x_t is the input to the ConvLSTM at time step t, h_{t-1} is the value of the hidden cell at the previous time step $t-1$, w represents the weights of the convolutional layer that convolves over the concatenation of x_t and h_{t-1}, and b represents the bias values used in each gate. The convolutions performed on the concatenation of x_t and h_{t-1} can be represented as gates, where $\mathbf{f_t}$ is the forget gate, $\mathbf{i_t}$ is the input gate, $\mathbf{o_t}$ is the output gate, and $\mathbf{g_t}$ is the control gate. The cell state at time step

t, c_t, is given by equation (3.7). The cell state represents the long term memory of the ConvLSTM model, which is passed through each of the LSTM cells in the model. The hidden state at time step t, h_t is given by equation (3.8), and represents the short term memory of the ConvLSTM model and can also be used as the models prediction at time step $t+1$. The equations above represent a single LSTM cell that is used in the ConvLSTM model, and an illustration of the process is depicted in Figure 3.2. Several LSTM cells can be stacked onto eachother, where the cell state and hidden state of the current LSTM cell is passed on to the next LSTM cell which handles the input at the next time step.

3.3.4 Focal Loss

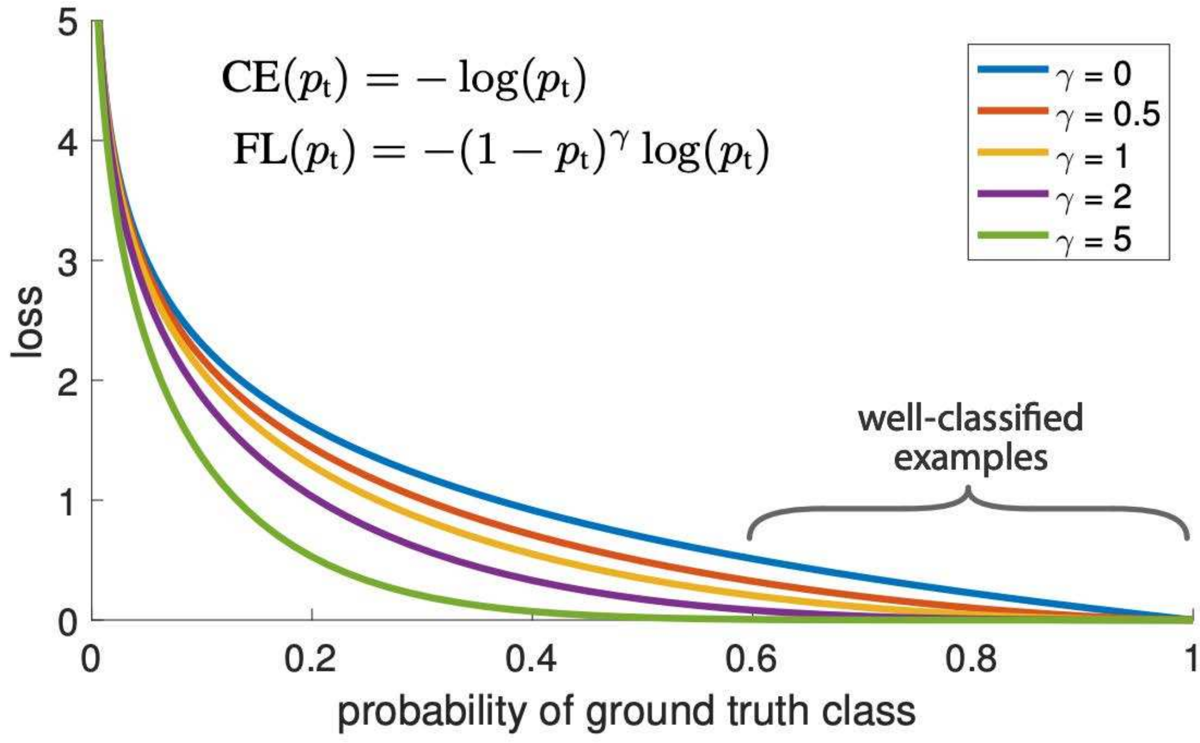

Figure 3.3: A description of the focal loss algorithm from [62]. The curves for focal loss are shown with different values of γ, which demonstrates how γ can control what samples are considered to be well-classified by the model.

Focal loss [62] is used as the loss function to train both the WPN and ConvLSTM models. The NEXTDAYWILDFIRESPREAD and FIRSTWEEKWILDFIRESPREAD datasets are heavily imbalanced, with most of the pixels belonging to the no fire class, and only a relative handful of pixels belonging to the fire class. Due to the larger image sizes in FIRSTWEEKWILDFIRESPREAD, this imbalance is also more pronounced than it is in NEXTDAYWILDFIRESPREAD. However, focal loss is able to handle class imbalanced datasets due to its ability to minimize the impact that the majority class has on the overall loss and ideally help the model focus more so on the minority class. Even if the model is able to classify no fire pixels easily, the small amount of loss incurred from each no fire pixel prediction can overwhelm the loss from the fire pixel predictions. The focal loss algorithm is given by the following equations,

$$FL(p_t) = -(1 - p_t)^{\gamma} \log(p_t), \tag{3.9}$$

$$p_t = \begin{cases} p, & \text{if } y = 1 \\ (1 - p), & \text{otherwise} \end{cases} . \tag{3.10}$$

Where p represents the model output for each pixel, which is a probability that a pixel is the fire class, and y is the class of the ground truth pixel (1 = fire, 0 = no fire). Because p is a probability, it is in the range of [0, 1]. The value p_t thus represents the probability for each of the models' outputs depending on what the ground truth pixel value y is. The γ value is the crucial component of the focal loss that essentially controls what is considered to be a well classified example. The higher the value of γ is, the further away the model's output needs to be from the ground truth probability to be considered well classified. In this context, a pixel that is well classified is one that contributes very little if anything to the total loss. Figure Figure 3.3 demonstrates the effect that γ has on the focal loss. We can also introduce a weighting factor α in the range of [0, 1], which can further help control the impact that the majority class has on the total loss. Then we can redefine the focal loss algorithm as

$$FL(p_t) = -\alpha_t(1 - p_t)^{\gamma} \log(p_t), \tag{3.11}$$

$$\alpha_t = \begin{cases} \alpha, & \text{if } y = 1 \\ (1-\alpha), & \text{otherwise} \end{cases}. \tag{3.12}$$

3.3.5 Model Training and Evaluation

The WPN models were trained with the RMSProp optimizer with a learning rate of 0.001 for 5 epochs each, while the ConvLSTM models were trained with the AdamW optimizer with a learning rate of 0.001 for 100 epochs. The ConvLSTM models utilized early stopping by ending training after 20 epochs of no improvement in the F1 score on the validation set. The WPN and ConvLSTM models utilized the focal loss, but with different parameters. For both the WPN and ConvLSTM models $\gamma = 2$ was used, but the weighting factor α was chosen to be 0.85 and 0.95 for the WPN and ConvLSTM models respectively. Both models were trained distributively on two RTX 3090's on a machine running Ubuntu 22.04 with 64 cores and 252GB of RAM. the WPN models used a batch size of 50, while the ConvLSTM models used a batch size of 100. The DQF, RGB, weather, elevation, and land cover bands were concatenated channel wise to form 11 channel images. The only augmentation performed on the data before it was passed to the models was a 200 x 200 center crop.

All of the models were evaluated on a test set which contained 8 fires that were not contained in the train and validation sets. The performance of the models was measured using the F1 score, which combines both precision and recall into a single metric, and the intersection of union (IOU). The IOU scores how well the fire pixels in the predicted fire masks overlap with the fire pixels in the ground truth fire masks. [20] uses a persistence algorithm on the NEXTDAYWILDFIREDATASET, which just takes the previous fire mask and uses that as the prediction. Following their lead, the F1 score of a persistence baseline is used to compare against the F1 scores of each model.

Chapter 4

Results

The results for the WPN model are summarized in Table 4.1, and Figure 4.1 compares how well the model performed against the persistence baseline. WPN outperformed the F1 score of the persistence baseline on only the 720 and 1440 minute ahead experiments. For reference, the WPN's performance on NEXTDAYWILDFIRESPREAD is also included in Table 4.1, and it appears to have performed worse on FIRSTWEEKWILDFIRESPREAD despite having a similar training process. However, there are several differences between the NEXTDAYWILDFIRESPREAD and FIRSTWEEKWILDFIRESPREAD datasets, which can be attributed to worse performance. Namely, there are different features being used and the characteristics of the fire products generated by the MODIS and ABI sensors are different to an extent. While the WPN model performed better when predicting at smaller time steps, it still did not outperform the persistence baseline which raises questions about its ability to perform tasks such as nowcasting for wildfire spreading. This underperformance could be attributed to a weak hyperparameter search, but the model itself may be insufficient for performing the task.

Table 4.1: A summary of the results from the WPN model predictions. Results of the WPN when using all 12 input features from NEXTDAYWILDFIRESPREAD are also included for comparison.

Mins. Ahead	Precision	Recall	F1	Persist. F1	Mean IOU	Max IOU
5	78.2%	83%	0.806	0.844	0.4565	1.0
30	61.9%	73.9%	0.674	0.718	0.3634	1.0
60	52.6%	65.9%	0.585	0.63	0.2873	1.0
120	44.9%	55.1%	0.495	0.514	0.202	1.0
180	36.1%	47.9%	0.412	0.428	0.1528	1.0
360	38.4%	14.4%	0.209	0.256	0.0313	0.7273
720	37%	14.1%	0.204	0.157	0.0199	1.0
1440	17.4%	36.5%	0.236	0.190	0.0457	0.875
1440 (NDWS)	30.3%	44.8%	0.361	0.309	0.1444	0.8947

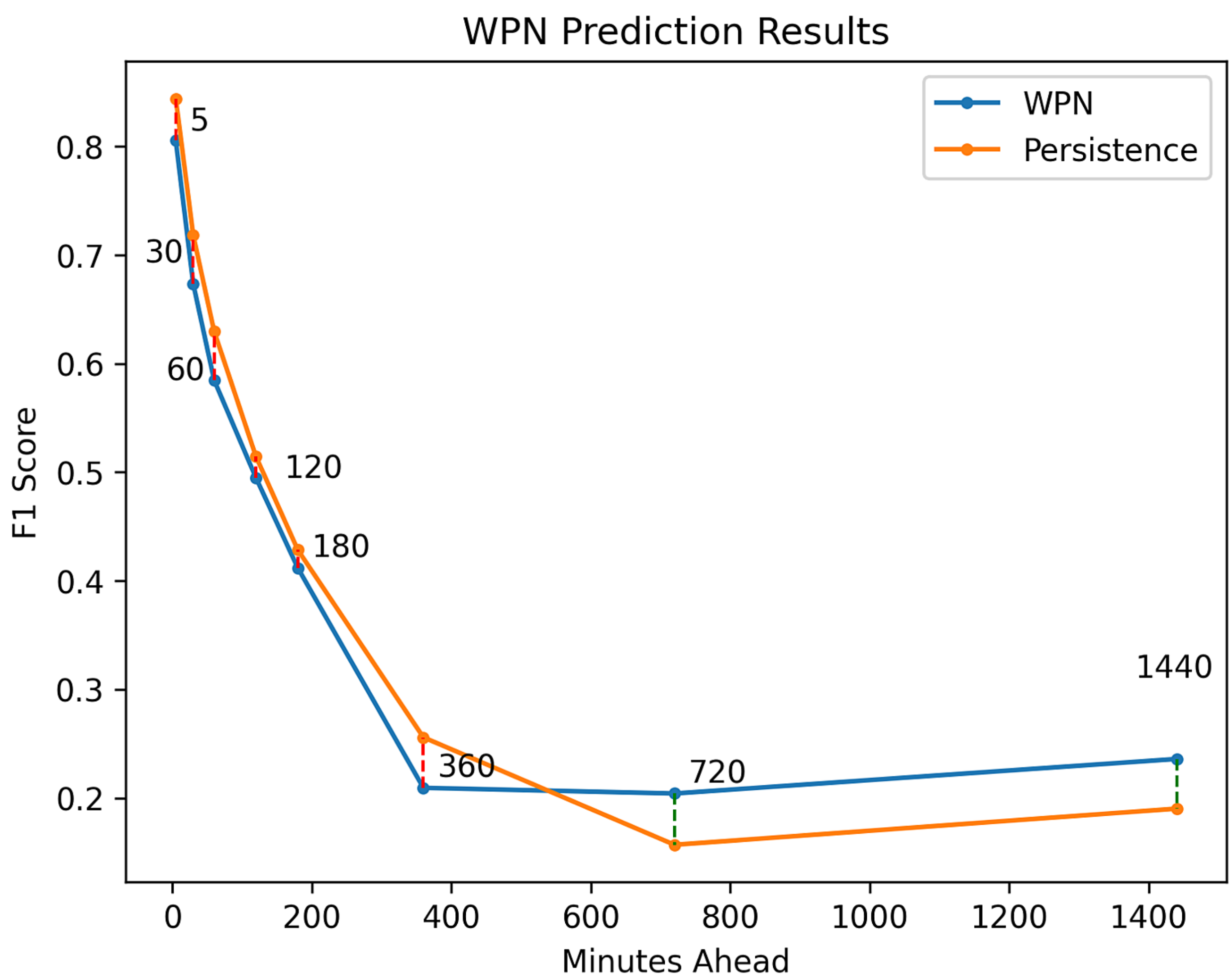

Figure 4.1: A plot showing the F1 scores from the predictions made by the WPN model. The WPN model was trained 8 separate times for 5 epochs on samples that were generated for predicting 5, 30, 60, 120, 180, 360, 720, and 1440 minutes ahead.

Test results from the ConvLSTM were unfortunately lackluster, with performances that appear to be worse than both the WPN and the persistence baselines. It's important to note though that there are less samples in the time series test sets used by the ConvLSTM model, especially when predicting at larger time steps. For example, the ConvLSTM model did slightly outperform the WPN model when predicting 360 minutes ahead, but there were only 13,260 samples compared to the 15,501 samples in the test set used by the WPN for 360 minutes ahead predictions. Interestingly, the ConvLSTM model had only 436 learnable weight parameters compared to the 8.7 million parameters in the WPN, which makes its performance seem slightly more impressive. Table 4.2 summarizes the results of the ConvLSTM, and Figure 4.2 compares the F1 scores of the ConvLSTM models to a persistence baseline.

Table 4.2: A summary of the results from the ConvLSTM model predictions.

Mins. Ahead	Precision	Recall	F1	Persist. F1	Mean IOU	Max IOU
5	72.8%	85.1%	0.785	0.844	0.4289	1.0
30	49.4%	62.3%	0.551	0.72	0.2229	1.0
60	32.7%	53.3%	0.405	0.632	0.1294	1.0
120	22.8%	46.3%	0.306	0.517	0.0910	0.6667
180	24.7%	29.2%	0.267	0.431	0.0939	1.0
360	20.4%	33.9%	0.255	0.275	0.0819	1.0
720	0.05%	3.1%	0.001	0.168	0.0005	0.0765
1440	0.03%	10.2%	0.001	0.18	0.0001	0.0189

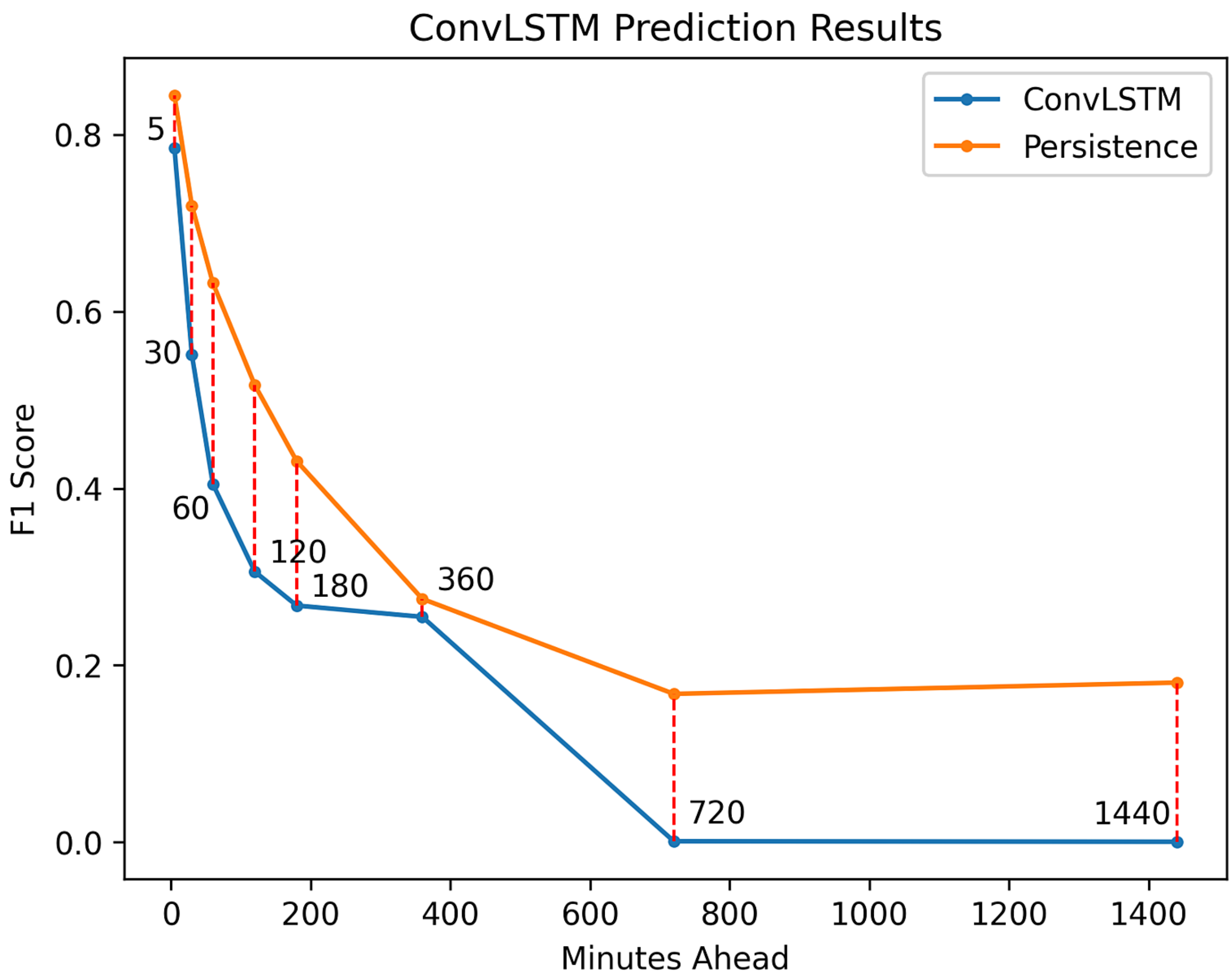

Figure 4.2: A plot showing the F1 scores from the predictions made by the ConvLSTM model. The ConvLSTM model was trained 8 separate times for ~100 epochs on samples that were generated for predicting 5, 30, 60, 120, 180, 360, 720, and 1440 minutes ahead. The ConvLSTM models used 5 previous updates on the fire locations at times t, $t-1$, $t-2$, $t-3$, and $t-4$.

Chapter 5

Discussion

There could be several reasons why the WPN model performed better on NEXTDAYWILDFIRESPREAD and worse on FIRSTWEEKWILDFIRESPREAD at the same time step of predicting 24 hours. This could be due to the differences in the Advanced Baseline Imager (ABI) sensor aboard GOES-16, and the MODIS sensor aboard Terra. ABI has a special resolution of 2km, while MODIS has a spatial resoultion of 1km. Furthermore, the quality of the MODIS data may be better, but no analysis has been done in this work to show that. It could also be the case that the training, testing, and validation sets from NEXTDAYWILDFIRESPREAD all contained fires that occurred in the same region, thus potentially giving the model an edge in the testing set due to the familiarity of the environment.

An interesting phenomenon that the occurs in wildfires is that they tend to die down at night when the temperatures are lower, and then ramp up during the day. This diurnal nature of wildfires can be clearly seen in the GOES images. This could be an explanation as to why the WPN model seemingly performed better on the 1440 minutes ahead experiments compared to the 360 and 720 minutes ahead experiments. When predicting 6 or 12 hours ahead, there are windows of time where the model needs to learn to predict that the fire will die down or pick up depending on the time of day. Information about the time of day is not explicitly fed to the models, but could potentially be inferred from some of the features such as RGB. The RGB channels have low pixel values at night which create dark images, and during the day there is more color. It's important to note though, the WPN model had higher precision when predicting 360 and 720 minutes ahead, but worse recall compared to the 1440 minutes ahead experiment.

5.1 Difficulties Encountered in GEE

Some difficulties were encountered when collecting images from the GEE datasets, especially the GOES-16 dataset. When making multiple requests to download an image for the same region and time, this would result in the possibility of receiving two different images. In other words, there were two different versions of the image that could be downloaded from the same URL. One of the versions tended to be downloaded more often, and the image that was less likely to be downloaded had a slight yet consistent offset both temporally and spatially. This led to the need to create a solution that would detect when the offset image was downloaded, and keep the non-offset image. However, this phenomenon seemed to occur sporadically, and there were days where downloading GOES-16 images did not result in the possibility of getting a slightly offset version.

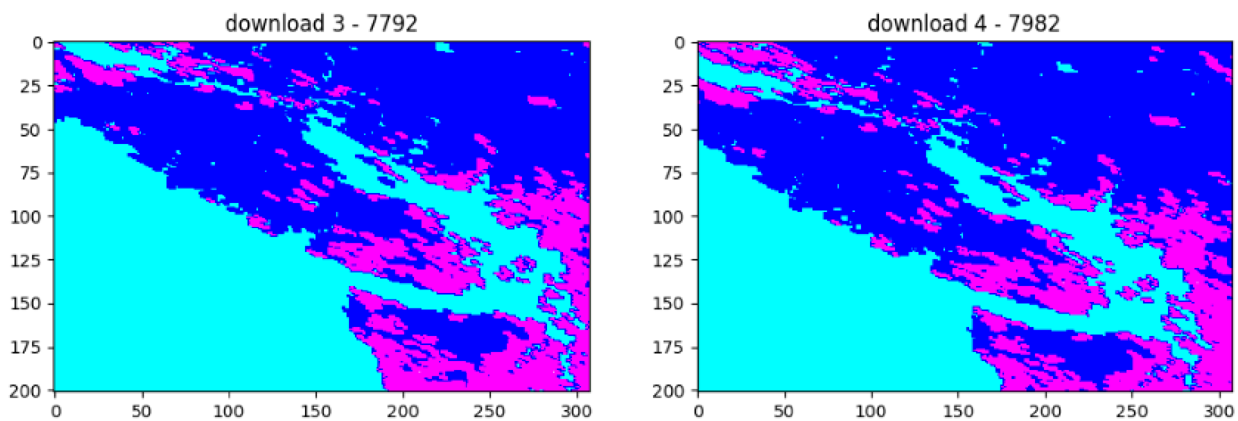

Figure 5.1: This shows two GOES-16 DQF images downloaded from the same Google Earth Engine URL. Both images are supposed to be of the same region and time, but they are slightly offset from one another. This offset can be seen by focusing on the top left corners of both of the images. They also have different download sizes in bytes, which is indicated in their titles.

GEE limits the total size of images that you can download, meaning you can't download images that have more than a certain amount of pixels. However, all image sizes that GEE lets you download are likely to be sufficient for performing wildfire spread prediction, and it's unclear yet if model performance improves as image sizes increase. It may be beneficial to have more context of weather patterns in the surrounding area of a fire, and even weather forecast data could prove to be useful. However, more information about other modalities, such as topography and fuel, that are further away from the fire locations are not likely to have significant impacts on the behavior of the fires. Including more pixels can also make it more difficult to train models for wildfire spread prediction, since most of the newly introduced pixels are likely to be in the no fire class. This will make it more difficult to control the contributions of the no fire pixels to the total loss and prevent them from overwhelming it.

5.2 Limitations

FIRSTWEEKWILDFIRESPREAD only takes the first week of megafires [63] into consideration. This criteria should be relaxed by reducing the total burned acres requirement to consider wildfires that have less than 100,000 total acres burned. Moreover, the structure of the dataset is less suitable for use as a benchmark. Unlike NEXTDAYWILDFIRESPREAD, which has a narrow focus on predicting 24 hours ahead, FIRSTWEEKWILDFIRESPREAD has a large variety of samples that can be created from it. It does not have predefined train, test, and validation splits that can be used by the community to report and compare results. Instead, users of FIRSTWEEKWILDFIRESPREAD are encouraged to be creative and come up with their own tasks and dataset splits that can be used to train a diverse range of models.

Leaving out certain features such as vegetation indices, energy release component (ERC), and population may have impacted performance. The assumption was that the land cover feature contains this information to an extent, but is it enough to completely replace the vegetation, ERC, and population features? Land cover has information about urban areas and vegetation, but it may be sacrificing more fine grained details in exchange for a more general overview of the landscape, and it could hurt model performance and analysis. Furthermore, there were no ablation studies performed in this work, which means that there are no insights into which features or components of the models are most important. Previous works have explored the importance of features [20, 32], but this is still an open question.

5.3 Future Works

Exploring the application of different models on the FIRSTWEEKWILDFIRESPREAD dataset such as transformers and generative AI could prove to be fruitful. Transformer models have been adapted for and are known to work well on vision related tasks [64], and there have also been transformer variants developed specifically for semantic segmentation [65, 66]. Generative models such as generative adversarial networks (GANs) [67] or variational autoencoders (VAEs) [68] could be good tools for providing synthetic wildfire spreading data. In addition, different ways of augmenting the data should be explored such as performing rotations and transformations, and determining how to appropriately add noise to the data can increase the robustness of models.

The FIRSTWEEKWILDFIRESPREAD dataset can be further expanded with more fires, and with more features (including those that were dropped from NEXTDAYWILDFIRESPREAD). It currently only includes massive fires with greater than 100,000 acres burned, which is only a small number of the total fires reported in the WFIGS IFP dataset. Including fires with less total acres burned can be more representative of common fires, and can help with understanding how useful machine learning models are when predicting more common wildfire spreading scenarios. Including only the first week of each fire was made with the intention that the data would have minimal ongoing containment efforts since it was early on into the fire. However, the models should be expected to make accurate predictions throughout the lifetime of a fire even while containment efforts are ongoing. Including all relevant data from the beginning to the end of a wildfire would be desirable for training and testing models on all the stages in the evolution of wildfires. Capturing the full lifetime of a wildfire could also help with investigating the spreading rate by pinpointing certain days where the conditions caused the fires to rapidly increase in size.

Chapter 6

Conclusion

In this work, the FIRSTWEEKWILDFIRESPREAD dataset was introduced which addresses gaps in existing datasets for wildfire spread prediction. FIRSTWEEKWILDFIRESPREAD contains remote sensing data with very high temporal resolutions, most notably utilizing the GOES-16 satellite to provide fire masks and RGB images with an impressive temporal resolution of 5 minutes. Metadata is also included with FIRSTWEEKWILDFIRESPREAD to improve replication efforts. A large variety of experiments can be performed with FIRSTWEEKWILDFIRESPREAD, samples can be created for predicting at virtually any multiple of 5 minutes ahead. Moreover, the dataset can be used to create time series samples which can be used to train a whole new category of machine learning models. The variety in experiments that can be performed with FIRSTWEEKWILDFIRESPREAD is demonstrated by training a U-Net based model that was originally designed for the Next Day Wildfire Spread dataset, and a convolutional LSTM model that can take advantage of the time series samples that FIRSTWEEKWILDFIRESPREAD can provide.

Bibliography

[1] National Interagency Fire Center. Federal firefighting costs (suppression only). Accessed: 9-30-2024. https://www.nifc.gov/fire-information/statistics/suppression-costs.

[2] Richard C Rothermel. *A mathematical model for predicting fire spread in wildland fuels*, volume 115. Intermountain Forest & Range Experiment Station, Forest Service, US ..., 1972.

[3] Mark A Finney. *FARSITE, Fire Area Simulator–model development and evaluation*. Number 4. US Department of Agriculture, Forest Service, Rocky Mountain Research Station, 1998.

[4] Joseph Quartieri, Nikos E Mastorakis, Gerardo Iannone, Claudio Guarnaccia, et al. A cellular automata model for fire spreading prediction. In *Latest Trends on Urban Planning and Transportation*, volume 1, pages 173–178. WORLD SCIENTIFIC AND ENGINEERING ACAD AND SOC, 2010.

[5] Giuseppe A Trunfio, Donato D'Ambrosio, Rocco Rongo, William Spataro, and Salvatore Di Gregorio. A new algorithm for simulating wildfire spread through cellular automata. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 22(1):1–26, 2011.

[6] Maria Vittoria Avolio, Salvatore Di Gregorio, Valeria Lupiano, and Giuseppe A Trunfio. Simulation of wildfire spread using cellular automata with randomized local sources. In *Cellular Automata: 10th International Conference on Cellular Automata for Research and Industry, ACRI 2012, Santorini Island, Greece, September 24-27, 2012. Proceedings 10*, pages 279–288. Springer, 2012.

[7] Lucia Russoa, Dimitris Vakalisb, and Constantinos Siettos. Simulating the wildfire in rhodes in 2008 with a cellular automata model. *CHEMICAL ENGINEERING*, 35, 2013.

[8] Andrés Cencerrado, Ana Cortés, and Tomàs Margalef. Genetic algorithm characterization for the quality assessment of forest fire spread prediction. *Procedia Computer Science*, 9:312–320, 2012.

[9] Andrés Cencerrado, Ana Cortés, and Tomàs Margalef. Applying probability theory for the quality assessment of a wildfire spread prediction framework based on genetic algorithms. *The Scientific World Journal*, 2013(1):728414, 2013.

[10] Jorge Pereira, Jérôme Mendes, Jorge SS Júnior, Carlos Viegas, and João Ruivo Paulo. A review of genetic algorithm approaches for wildfire spread prediction calibration. *Mathematics*, 10(3):300, 2022.

[11] Baisravan HomChaudhuri, Manish Kumar, and Kelly Cohen. Genetic algorithm based simulation–optimization for fighting wildfires. *International Journal of Computational Methods*, 10(06):1350035, 2013.

[12] Virginia G Gasull, DF Larios, Julio Barbancho, Carlos León, and Mohammad S Obaidat. Computational intelligence applied to wildfire prediction using wireless sensor networks. In *Proceedings of the International Conference on*, pages 1–8. IEEE, 2011.

[13] Virginia G Gasull, DF Larios, Julio Barbancho, Carlos León, and Mohammad S Obaidat. A wildfire prediction based on fuzzy inference system for wireless sensor networks. In *E-Business and Telecommunications: International Joint Conference, ICETE 2011, Seville, Spain, July 18-21, 2011, Revised Selected Papers*, pages 43–59. Springer, 2012.

[14] Yen Kheng Tan and Sanjib Kumar Panda. Self-autonomous wireless sensor nodes with wind energy harvesting for remote sensing of wind-driven wildfire spread. *IEEE Transactions on Instrumentation and Measurement*, 60(4):1367–1377, 2011.

[15] 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.

[16] George E Sakr, Imad H Elhajj, George Mitri, and Uchechukwu C Wejinya. Artificial intelligence for forest fire prediction. In *2010 IEEE/ASME international conference on advanced intelligent mechatronics*, pages 1311–1316. IEEE, 2010.

[17] Youssef Safi and Abdelaziz Bouroumi. Prediction of forest fires using artificial neural networks. *Applied Mathematical Sciences*, 7(6):271–286, 2013.

[18] Xin Cao, Xihong Cui, Miao Yue, Jin Chen, Hiroki Tanikawa, and Yu Ye. Evaluation of wildfire propagation susceptibility in grasslands using burned areas and multivariate logistic regression. *International Journal of Remote Sensing*, 34(19):6679–6700, 2013.

[19] Sandra Oliveira, Friderike Oehler, Jesús San-Miguel-Ayanz, Andrea Camia, and José MC Pereira. Modeling spatial patterns of fire occurrence in mediterranean europe using multiple regression and random forest. *Forest Ecology and Management*, 275:117–129, 2012.

[20] 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.

[21] 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.

[22] Spyridon Kondylatos, Ioannis Prapas, Gustau Camps-Valls, and Ioannis Papoutsis. Mesogeos: A multi-purpose dataset for data-driven wildfire modeling in the mediterranean. *Advances in Neural Information Processing Systems*, 36, 2024.

[23] Dimitris Sykas, Dimitrios Zografakis, Konstantinos Demestichas, Constantina Costopoulou, and Pavlos Kosmidis. Eo4wildfires: an earth observation multi-sensor, time-series machine-learning-ready benchmark dataset for wildfire impact prediction. In *Ninth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2023)*, volume 12786, pages 11–20. SPIE, 2023.

[24] David Radke, Anna Hessler, and Dan Ellsworth. Firecast: Leveraging deep learning to predict wildfire spread. In *IJCAI*, pages 4575–4581, 2019.

[25] Wenyu Jiang, Yuming Qiao, Guofeng Su, Xin Li, Qingxiang Meng, Hongying Wu, Wei Quan, Jing Wang, and Fei Wang. Wfnet: A hierarchical convolutional neural network for wildfire spread prediction. *Environmental Modelling & Software*, 170:105841, 2023.

[26] Jonathan L Hodges and Brian Y Lattimer. Wildland fire spread modeling using convolutional neural networks. *Fire technology*, 55:2115–2142, 2019.

[27] 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.

[28] 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.

[29] Frédéric Allaire, Vivien Mallet, and Jean-Baptiste Filippi. Emulation of wildland fire spread simulation using deep learning. *Neural networks*, 141:184–198, 2021.

[30] Dmitrii Shadrin, Svetlana Illarionova, Fedor Gubanov, Ksenia Evteeva, Maksim Mironenko, Ivan Levchunets, Roman Belousov, and Evgeny Burnaev. Wildfire spreading prediction using multimodal data and deep neural network approach. *Scientific Reports*, 14(1):2606, 2024.

[31] Mohammad Marjani, Seyed Ali Ahmadi, and Masoud Mahdianpari. Firepred: A hybrid multi-temporal convolutional neural network model for wildfire spread prediction. *Ecological Informatics*, 78:102282, 2023.

[32] 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.

[33] Mohammad Marjani, Masoud Mahdianpari, Seyed Ali Ahmadi, Emadoddin Hemmati, Fariba Mohammadimanesh, and Mohammad Saadi Mesgari. Application of explainable artificial intelligence in predicting wildfire spread: an aspp-enabled cnn approach. *IEEE Geoscience and Remote Sensing Letters*, 2024.

[34] Louis Giglio, Wilfrid Schroeder, and Christopher O Justice. The collection 6 modis active fire detection algorithm and fire products. *Remote sensing of environment*, 178:31–41, 2016.

[35] John T Abatzoglou. Development of gridded surface meteorological data for ecological applications and modelling. *International journal of climatology*, 33(1):121–131, 2013.

[36] Andy Jarvis, Hannes I Reuter, Andy Nelson, Edward Guevara, et al. Hole-filled srtm for the globe version 4, available from the cgiar-csi srtm 90m database. 2008.

[37] A Didan and Armando Barreto. Viirs/npp vegetation indices 16-day l3 global 500m sin grid v001. *NASA EOSDIS Land Processes DAAC: Oak Ridge, TN, USA*, 2018.

[38] Center for International Earth Science Information Network CIESIN Columbia University. Gridded population of the world, version 4 (gpwv4): Population density, revision 11. Accessed: 21-9-2024.

[39] Tomàs Artés, Duarte Oom, Daniele De Rigo, Tracy Houston Durrant, Pieralberto Maianti, Giorgio Libertà, and Jesús San-Miguel-Ayanz. A global wildfire dataset for the analysis of fire regimes and fire behaviour. *Scientific data*, 6(1):296, 2019.

[40] Wilfrid Schroeder, Patricia Oliva, Louis Giglio, and Ivan A Csiszar. The new viirs 375 m active fire detection data product: Algorithm description and initial assessment. *Remote Sensing of Environment*, 143:85–96, 2014.

[41] Eric Vermote, Belen Franch, and Martin Claverie. Viirs/npp surface reflectance daily l2g global 1km and 500m sin grid v001, 2016.

[42] Shepard A Clough, Mark W Shephard, Eli J Mlawer, JS Delamere, Michael J Iacono, Karen Cady-Pereira, S Boukabara, and Patrick D Brown. Atmospheric radiative transfer modeling: A summary of the aer codes. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 91(2):233–244, 2005.

[43] Damien Sulla-Menashe, Josh M Gray, S Parker Abercrombie, and Mark A Friedl. Hierarchical mapping of annual global land cover 2001 to present: The modis collection 6 land cover product. *Remote sensing of environment*, 222:183–194, 2019.

[44] Dimitris Sykas, Dimitrios Zografakis, and Konstantinos Demestichas. Deep learning approaches for wildfire severity prediction: A comparative study of image segmentation networks and visual transformers on the eo4wildfires dataset. Accessed: 21-9-2024, 2024.

[45] Timothy J Schmit, Mathew M Gunshor, W Paul Menzel, James J Gurka, Jun Li, and A Scott Bachmeier. Introducing the next-generation advanced baseline imager on goes-r. *Bulletin of the American Meteorological Society*, 86(8):1079–1096, 2005.

[46] Timothy J Schmit, Paul Griffith, Mathew M Gunshor, Jaime M Daniels, Steven J Goodman, and William J Lebair. A closer look at the abi on the goes-r series. *Bulletin of the American Meteorological Society*, 98(4):681–698, 2017.

[47] C.O Justice, L Giglio, S Korontzi, J Owens, J.T Morisette, D Roy, J Descloitres, S Alleaume, F Petitcolin, and Y Kaufman. The modis fire products. *Remote Sensing of Environment*, 83(1):244–262, 2002. The Moderate Resolution Imaging Spectroradiometer (MODIS): a new generation of Land Surface Monitoring.

[48] Ivan Csiszar, Wilfrid Schroeder, Louis Giglio, Evan Ellicott, Krishna P Vadrevu, Christopher O Justice, and Brad Wind. Active fires from the suomi npp visible infrared imaging radiometer suite: Product status and first evaluation results. *Journal of Geophysical Research: Atmospheres*, 119(2):803–816, 2014.

[49] MK Bah, MM Gunshor, and TJ Schmit. Generation of goes-16 true color imagery without a green band. *Earth and Space Science*, 5(9):549–558, 2018.

[50] 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.

[51] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, et al. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*, 2018.

[52] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. Convolutional lstm network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, 28, 2015.

[53] Bernardt Duvenhage and Jan Jacobus Nel. Inaccuracies when mixing coordinate reference frameworks in a system of systems simulation. In *AFRICON 2007*, pages 1–8. IEEE, 2007.

[54] Volker Janssen. Understanding coordinate reference systems, datums and transformations. *International Journal of Geoinformatics*, 2009.

[55] Gabriela Csurka, Riccardo Volpi, Boris Chidlovskii, et al. Semantic image segmentation: Two decades of research. *Foundations and Trends® in Computer Graphics and Vision*, 14(1-2):1–162, 2022.

[56] Kanti V Mardia and Peter E Jupp. *Directional statistics*. John Wiley & Sons, 2009.

[57] Philipp Berens. Circstat: a matlab toolbox for circular statistics. *Journal of statistical software*, 31:1–21, 2009.

[58] Md Zahangir Alom, Mahmudul Hasan, Chris Yakopcic, Tarek M Taha, and Vijayan K Asari. Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. *arXiv preprint arXiv:1802.06955*, 2018.

[59] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19*, pages 424–432. Springer, 2016.

[60] Nahian Siddique, Sidike Paheding, Colin P Elkin, and Vijay Devabhaktuni. U-net and its variants for medical image segmentation: A review of theory and applications. *IEEE access*, 9:82031–82057, 2021.

[61] S Hochreiter. Long short-term memory. *Neural Computation MIT-Press*, 1997.

[62] T Lin. Focal loss for dense object detection. *arXiv preprint arXiv:1708.02002*, 2017.

[63] Grant D Linley, Chris J Jolly, Tim S Doherty, William L Geary, Dolors Armenteras, Claire M Belcher, Rebecca Bliege Bird, Andrea Duane, Michael-Shawn Fletcher, Melisa A Giorgis, et al. What do you mean,'megafire'? *Global Ecology and Biogeography*, 31(10):1906–1922, 2022.

[64] Alexey Dosovitskiy. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

[65] Robin Strudel, Ricardo Garcia, Ivan Laptev, and Cordelia Schmid. Segmenter: Transformer for semantic segmentation. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 7262–7272, 2021.

[66] Enze Xie, Wenhai Wang, Zhiding Yu, Anima Anandkumar, Jose M Alvarez, and Ping Luo. Segformer: Simple and efficient design for semantic segmentation with transformers. *Advances in neural information processing systems*, 34:12077–12090, 2021.

[67] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

[68] Diederik P Kingma. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.