Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19)

FireCast: Leveraging Deep Learning to Predict Wildfire Spread

David Radke[1*†], **Anna Hessler**[2] and **Dan Ellsworth**[2]
[1]David R. Cheriton School of Computer Science, University of Waterloo
[2]Department of Mathematics and Computer Science, Colorado College
dtradke@uwaterloo.ca, {a_hessler, dellsworth}@coloradocollege.edu

Abstract

Destructive wildfires result in billions of dollars in damage each year and are expected to increase in frequency, duration, and severity due to climate change. The current state-of-the-art wildfire spread models rely on mathematical growth predictions and physics-based models, which are difficult and computationally expensive to run. We present and evaluate a novel system, FireCast. FireCast combines artificial intelligence (AI) techniques with data collection strategies from geographic information systems (GIS). FireCast predicts which areas surrounding a burning wildfire have high-risk of near-future wildfire spread, based on historical fire data and using modest computational resources. FireCast is compared to a random prediction model and a commonly used wildfire spread model, Farsite, outperforming both with respect to total accuracy, recall, and F-score.

1 Introduction

Each year, wildfires destroy billions of dollars worth of infrastructure and claim the lives of people caught in their path [Finney *et al.*,]. Studies indicate climate change will cause more severe drought in vulnerable areas, leading to more dry and dead vegetation, increasing the duration and frequency of wildfires [Stephens *et al.*, 2018; Westerling *et al.*, 2006]. The United States Fourth National Climate Assessment Volume II highlights the dramatic expected future impact of climate change, including an increase in wildfire frequency and duration that is expected to impact the global economy, human health, infrastructure, and agriculture [Reidmiller *et al.*, 2018]. The Dynamic Integrated Model of Climate and the Economy (DICE) concludes that the total discounted economic damages of climate change with no abatement are in the order of $23 trillion USD in economic losses over the next 80 years [Barker, 2008]. With growing concerns about wildfire duration and frequency, increasing the efficiency of wildfire fighting is increasingly important.

Currently, most utilized fire modeling technologies are either solely informative of past fire perimeters or rely on mathematical growth predictions and physics-based models to predict how a fire will spread. These popular models rely on data collected during controlled forest burns, lab-burn experiments of vegetation, and physics-based simulation, or require constant and expensive drone data collection of the area to determine fire growth in a region [Finney, 1998; Forghani *et al.*, 2007; Lin *et al.*, 2019]. These systems do not always incorporate data from past wildfire events or measurements that are observed in nature from uncontrolled fire activity, and do not systematically learn from historic events. Although physics-based systems incorporate the impact of a large number of precise variables, gathering the necessary data for an active fire is often extremely difficult. The existing commonly used models are also computationally expensive to run, are generally focused on hourly activity of a fire, and most do not implement strategies and technology from artificial intelligence (AI), such as neural networks, to assist in the prediction process [Subramanian and Crowley, 2018].

We suggest that the problem of predicting wildfire growth is too complex for a single discipline to effectively mediate. We present FireCast, a novel solution that combines AI and Geographic Information Systems (GIS) to predict future wildfire spread, given a small number of location characteristics and a weather forecast. AI techniques have the ability to make classifications or predictions about a given target based on a set of input features, and GIS has the ability to generate the appropriate geospatial input variables for such an AI model. FireCast uses supervised learning and geospatial inputs, such as satellite imagery, elevation data, weather data, and historical fire perimeters to identify patterns associated with fire spread in certain environments to produce predictions of wildfire spread. The historical fire perimeters were manually mapped by fire fighters working to contain the fires. To our knowledge, FireCast is the first application of supervised machine learning for wildfire spread prediction.

While our training and evaluation data are limited to the Rocky Mountain region of the United States, we believe FireCast has the ability to scale to any region with proper training data. We conduct experiments and evaluate the predictive power of FireCast both statistically and visually, and compare against the most common predictive modeling software used by fire fighters today. The major contributions of this pa-

[*]Contact Author
[†]Started work as an undergraduate student at Colorado College.

per are: i) the presentation of a novel and lightweight system for predicting wildfire spread to support informed decision-making by experienced fire fighters; and ii) a comparison of AI's effectiveness at predicting wildfire growth. This work could help reduce the impact of wildfires, saving billions of dollars of infrastructure, and more importantly, saving lives.

2 Related Work

For decades, modeling and understanding the affects of wildfires has been a major area of research across multiple disciplines [Finney, 1994; Finney, 1998; Forghani *et al.*, 2007; Houtman *et al.*, 2013; Lauer *et al.*, 2017; Lin *et al.*, 2019; Radke, 1995; Stephens *et al.*, 2018; Séro-Guillaume *et al.*, 2008; Zhang *et al.*, 2016]. This section highlights some popular and recent efforts to solve similar problems.

2.1 Wildland Fire Decision Support System

The Wildland Fire Decision Support System (WFDSS) is a web-based geospatial fire management portal used by some state and federal fire agencies to manage and document large fires [O'Connor *et al.*, 2016]. WFDSS uses fire perimeter mapping and a modeling framework to record fire progression each day while incorporating real-time short-term weather forecasts, however WFDSS does not attempt to predict future growth. By avoiding predictions and only presenting the user with known information, WFDSS is designed for the sole purpose of supporting flexible and informed decision making by fire managers about the current condition of a fire.

2.2 Farsite

Finney, a Research Scientist with the U.S. Forest Service, is the lead developer of the most commonly used fire modeling tool, Farsite, which predicts fire perimeters using a two-dimensional (2D) deterministic fire growth model [Finney, 1994; Finney, 1998]. Landscape and vegetation characteristics along with weather data with physics and mathematical-based models are used to generate predicted future fire perimeters for a determined time period. The tool requires a large number of location based variables, such as a manually generated landscape file[1], weather, winds, fuels, fuel moisture, fire spread rate adjustments, along with other optional inputs. These files can be difficult and often expensive to generate for every location, often requiring surveyors on the ground to collect precise information.

The fundamental fire growth shape in Farsite is based on observations of past research which suggests 2D fire shapes are ellipsoidal, in theory, under uniform conditions. Pure ellipsoidal behavior is observed in uniform environments when factors affecting fire behaviors, such as fuels, weather and topography, are constant, which is uncommon in nature. Thus, the model incorporates mathematical interpretations of affecting conditions to shape the predicted fire perimeters, and altering a variable can dramatically change the predicted output.

[1]A landscape file is a manually generated file (.lcp) which contains location data for elevation, slope, aspect, fuel model, and canopy cover of vegetation.

2.3 Agent Based Wildfire Spread Prediction

Recent work has applied techniques from AI to predict spatial spreading of wildfires in a model called MCTS-A3C [Subramanian and Crowley, 2018]. MCTS-A3C uses a deep reinforcement learning approach in which the AI agent is the fire, given the task of spreading across the surrounding landscape. Similar to the approach of other models [Finney, 1998; Kourtz and O'Regan, 1971], the agent starts in one location and spreads outward along a satellite image.

MCTS-A3C uses a Markov Decision Process (MDP) to describe the state of any location on the landscape pertaining to temperature, land cover type, wind speed, wind direction, humidity, fire intensity, number of days since the start of the fire, and average amount of rainfall during the coarse of the study. MCTS-A3C uses a Monte Carlo Tree Search (MCTS) in which each node is a cell on fire and has its own state at that time [Kocsis and Szepesvári, 2006]. The fire is made to start at the ignition points of the tree, and each node of the search tree is comprised of a cell burning in the satellite image. The model uses data from simulated and historic forest fire events as validation data for their model. The ground truth data resolution used as input to MCTS-A3C varies from 30 meters to 300 meters for some locations. After various experiments, the authors conclude that MCTS-A3C has a high burn accuracy for all domains and environments tested.

3 FireCast Algorithm

FireCast is different from past wildfire spread prediction research because of the incorporation of deep supervised machine learning methods in a unique model structure. Recently, Convolutional Neural Networks (CNN) have been used with remotely sensed data for different tasks, and show powerful capabilities [Ding *et al.*, 2018; Maggiori *et al.*, 2017; Zhang *et al.*, 2016]. Therefore, we implement the 2D CNN detailed in Table 1, which is used for supervised learning from the various visual inputs explained further in §4, and is implemented using Keras with the TensorFlow backend.

The full CNN is composed of two convolutional layers of 32 and 64 hidden nodes, and uses Sigmoid and ReLU activation functions respectively. A sliding window explores the square area of 30 pixels around each sampled pixel-of-interest (POI) of a given visual layer, with an actual kernel size of 3×3. Additionally, the CNN has one average pooling layer, two max pooling layers, and three dropout layers. The CNN output tensor is concatenated with the location's eight atmospheric data points and used as input to a dense layer using a Sigmoid activation function, which maps to a single output value for each pixel. Both the CNN and the final dense layer use the binary crossentropy loss function and do not use regularizers. The CNN uses an RMSProp optimizer and the final dense layer uses a stochastic gradient decent optimizer.

FireCast is trained to predict the areas surrounding the current fire perimeter that are expected to burn during the following 24 hours given an initial fire perimeter, location characteristics, and atmospheric data as input.

The final model is tested on a historical fire with consecutively mapped perimeters, referred to as the *testing fire*, that was omitted from training. FireCast receives the same input

Layer	Operation	Kernel/Pool Size	Feature Maps
1	Avg Pooling	2×2	–
2	Convolution	3×3	32
	Max Pooling	2×2	–
	Dropout	–	–
3	Convolution	3×3	64
	Max Pooling	2×2	–
	Dropout	–	–
4-Out	Dense	–	128
5	Concat	–	136
6-Out	Dense	–	1

Table 1: The FireCast model architecture.

variables as the training fires for the testing fire, starting with the first recorded fire perimeter and making next day predictions for all collected perimeters. FireCast randomly samples POIs surrounding the current fire perimeter and assigns a prediction value p, so that $p \in [0, 1]$, to each pixel representing the likelihood the model predicts the pixel will be contained within the following fire perimeter given the input variables. Higher values represent a higher predicted likelihood of burn.

4 FireCast Input Data

The ground truth visual and atmospheric training data are gathered from consecutively mapped days of historical fires. While it is appropriate to collect visual inputs such as physical geography and satellite imagery a single time just before the fire ignition, other information such as fire perimeters and atmospheric data are gathered for regular 24-hour and 1-hour intervals respectively. Due to the small number of mapped perimeters, we augment the data collected for training to generate a larger training dataset [Simard *et al.*, 2003]. For each fire in the training dataset, referred to as *training fires*, the model is exposed to an initial fire perimeter and the future 24 hours of weather data. The ground truth fire perimeter for the following day is used as validation data during training.

Our evaluation uses historical fire perimeters from GeoMAC, a United States Geological Survey (USGS) database, for each of the training and testing fires[2]. The GeoMAC database contains perimeter data for a variety of fires, although the majority of these fires do not contain consecutively mapped perimeters in 24 hour intervals. Consecutive days are targeted for collection since they best support the goal of next-day predictions.

Landsat8 satellite imagery is used as a visual input to the model, collected from GloVis[3], another USGS database. Landsat8 is collected on an interval of every few months, so the selected images are the most recent cloudless Landsat8 image of each location prior to the ignition of the fire. The imagery has a resolution of 30 meters, where each pixel represents a 30×30 meter square on the ground. FireCast only uses the red, blue, green, and near-infrared bands since those features are known to correlate with vegetation on the ground and support abstractions such as a Normalized Difference Vegetation Index (NDVI) for normalized vegetation health levels so that $NDVI \in [-1, 1]$ for each pixel, describing the health of the vegetation in that pixel, where a greater value describes healthier vegetation [Carlson and Ripley, 1997].

We also collect a digital elevation model (DEM), where each pixel is the elevation value at that particular point. A 30-meter resolution DEM is obtained from the USGS National Map[4] for each fire location and geo-rectified with the Landsat8 pixels. The DEM is further processed to derive landscape aspect, where each pixel contains a degree from north for the direction the ground is facing, so that degree $\in [0, 359]$.

Weather has a significant impact on any geographically based model, especially when predicting the highly variable phenomena of wildfire growth. Therefore, the most precise and accurate historical atmospheric data available for each remote location was collected from the National Oceanic and Atmospheric Administration (NOAA)[5], including atmospheric pressure, temperature (Celsius), dew point, wind direction, wind speed, precipitation, and relative humidity for each fire location. Each burn day from the training and testing fires is exposed to 24 "future" hours of atmospheric data to accompany the perimeter at midnight of that day. In a real fire situation, hourly predictions of forecasts are used in place of this historical input data. The most precise atmospheric data available from NOAA is collected on a three-kilometer scale, hence, the weather is interpolated for each region to match the 30 m resolution of the other inputs.

5 FireCast Output

The output from FireCast for each day of the testing fire is an image of the area that displays the sampled POIs which are colored corresponding to their predicted values to burn. The model uses all of the visual input layers, a starting known fire perimeter, and the next 24 hours of atmospheric data for that location, to generate the output image that displays areas that the model believes are locations of high risk for fire spread. If there are multiple days of input for the testing fire, the algorithm will produce the corresponding number of unique output images. The POIs are randomly sampled from pixels outside of the initial fire perimeter. The model assigns a value to each POI based on the its characteristics and the surrounding pixels in the sliding window.

6 Evaluation

In this paper, FireCast is statistically evaluated for total prediction accuracy, total burn accuracy, and F-score. In addition, predicted FireCast outputs are visually compared to actual perimeters of the test fire, the 2016 Beaver Creek, Colorado fire. For the duration of the evaluation process, the total set of predictions, T_p, is classified by:

$$T_p = \begin{cases} 1 & \text{if } p \geq 0.5 \\ 0 & \text{if } p < 0.5 \end{cases} \quad (1)$$

[2]https://www.geomac.gov

[3]https://glovis.usgs.gov/app?fullscreen=0

[4]https://www.usgs.gov/core-science-systems/national-geospatial-program/national-map

[5]https://www.ncdc.noaa.gov/cdo-web/

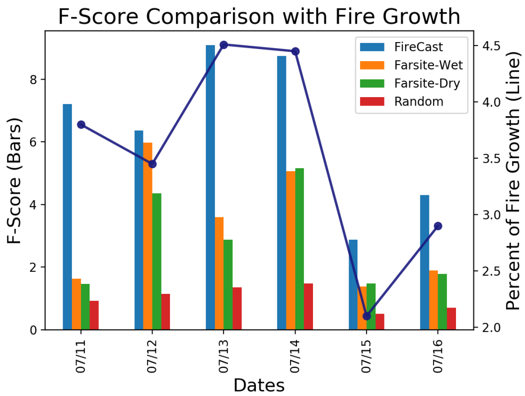

Figure 1: Comparing F-scores of FireCast, Farsite, and a random model. The line represents the percent of fire growth.

where 1 represents a predicted burn and 0 represents a predicted non-burn. To show that FireCast actually learns fire patterns despite limited training data, we compare statistically against random pixel predictions and Farsite, the most widely used fire prediction software today.

Unlike Farsite, fuel moisture is not required for FireCast and is not available for the testing fire location. Therefore, Farsite is run twice for each input perimeter with wetter and drier fuel moisture files, believing that the ground truth falls within those parameters. Farsite is initialized to burn from an initial perimeter with known atmospheric data, and computes a growth prediction 24 hours into the future.

6.1 Total Prediction Accuracy

Total prediction accuracy, or how many of the randomly sampled pixels are correctly predicted, is one method of evaluating a pixel-based prediction model. Let $SetCorrect_{Dp}$ denote the set of predictions p made that are correctly classified as either burned or not burned for day D; and let $TotalPredictions_{Dp}$ represent all of the predictions p made for the randomly sampled pixels during D. Thus, total prediction accuracy is found for each day by,

$$Acc_{Dp} = \frac{|SetCorrect_{Dp}|}{|TotalPredictions_{Dp}|}. \quad (2)$$

A higher accuracy expresses greater predictive classification capabilities of the model. Over all days of the testing fire, we observe an average accuracy of 87.7% for FireCast, 50.4% for the random predictions, 67.8% and 63.6% for Farsite with wet and dry fuel moisture respectively.

This metric alone is not an effective methodology to test for overall performance. For example, one could attain high accuracy by including more land further from the fire, lowering the ratio of pixels highly vulnerable to burn in the whole area. In this scenario, the model is less likely to predict the pixels far away from the fire will burn, thus increasing the overall prediction accuracy.

6.2 Recall

We define recall as the fraction of correctly classified pixels where $p \geq 0.5$ over the actual ground truth amount of

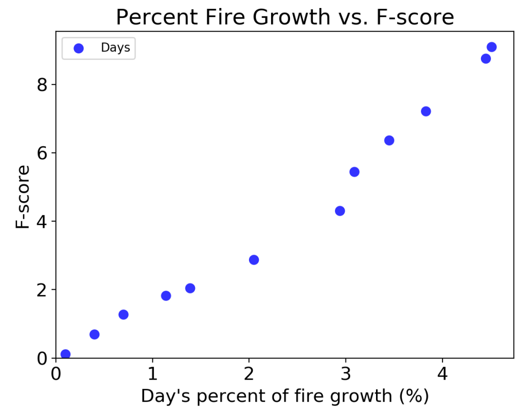

Figure 2: Comparisons between the F-score and percent of new burn for two chunks of consecutively mapped days of the testing fire.

the sampled pixels that burned in the following 24 hours. In theory with respect to fire modeling, we suggest it is better to over-predict areas than to under-predict, especially where neighborhoods are possibly at risk. Let $PredictedBurn_D$ denote the set of pixels that FireCast predicts will burn in the next day D, and $TotalBurn_D$ be the set of all sampled pixels on day D that actually did burn. Thus, recall is calculated as,

$$Recall_{Dp} = \frac{|PredictedBurn_D|}{|TotalBurn_D|}. \quad (3)$$

For the testing fire, we observe an average recall of 91.1% for FireCast, 50.4% for the randomly assigned pixels, 74.8% and 81.1% for wet and dry condition Farsite models respectively. Thus, FireCast outperforms current fire modeling technologies with respect to recall, although similar to total accuracy, it alone is not an effective evaluation methodology; a model where $p \geq 0.5$ for all p will have a recall of 100%.

6.3 F-Score

The F-score is a statistical method of evaluation, defined as the harmonic average of both precision and recall. We define precision as the fraction of correctly classified burn pixels where $p \geq 0.5$ over the total amount of pixels that FireCast predicts to burn in the following of 24 hours.

A higher F-score value means the model is more effective at overall classification of the sampled pixels for prediction. On average, FireCast achieves an average precision of 3.6% over six consecutive days of the testing fire, yielding an average F-score of 6.4%. For the same days, the randomly predicted pixels produce an average F-score of 1.0%, and Farsite produces 3.5% and 2.9% for wet and dry conditions respectively. Therefore we conclude that FireCast outperforms the random prediction model and Farsite when evaluated by the F-score, detailed in Figure 1. This comparison shows that FireCast does learn patterns of fires spread from different landscapes in the training data, and is applicable to a new landscape to outperform current widely used software in a lightweight and less computationally expensive solution.

We note FireCast typically has a higher F-score on days with more fire growth. Figure 2 visualizes this upward trend

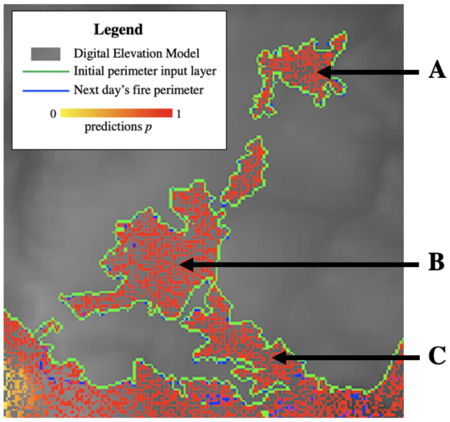

Figure 3: Visualization of fire islands and a fire valley.

by investigating the relationship between the percentage of fire growth and the F-score for each day of the testing fire, which is consistent with Figure 1. Due to limited recorded perimeter data, our testing fire dataset observes a two week gap in data from July 16, 2016 to August 2, 2016. We compare the pixel prediction output from July 16th with the next recorded ground truth perimeter on August 2nd and evaluate the pixel predictions. Over this period, the testing fire grew a massive 52.6%, and the comparison produces an F-score of 34.4% for FireCast. The trend in Figure 2 is consistent over the two week gap, showing that FireCast has potential predictive power at least two weeks into the future, whereas Farsite has an F-score of 1.3% and 1.2% for the wet and dry scenarios respectively for the same comparison. We suggest the reason behind this is both unreported fire fighter activity, and a spotting phenomena that we further explain in §6.5.

6.4 FireCast Visual Output Legend

For each FireCast output figure presented in this paper, the features follow the same legend as in Figure 3. The background of the output is the DEM input layer to provide the user with a general understanding of the geography of the location. The initial perimeter input layer is displayed as a green outline, and the next day's fire perimeter is displayed for the user as a blue perimeter outline, though this perimeter is hidden from the model. The sampled pixels for prediction are colored on a scale from yellow to red. Yellow represents pixels where p is closer to 0 and red where p is closer to 1.

6.5 Visual Evaluation

FireCast is designed to support informed decision making by first responders, thus a visual evaluation of the output is important and useful when determining the validity and performance of the model. The visual evaluation of the output suggests that the biggest advantage of FireCast is its ability predict high-risk areas of large and rapid fire spread, and presents the notion that the pixels incorrectly predicted to burn can actually be valuable for strategic purposes. In general, the areas in the output image with highly dense clusters of pixels incorrectly predicted to burn either have unique physical characteristics, or were eventually overtaken by the fire.

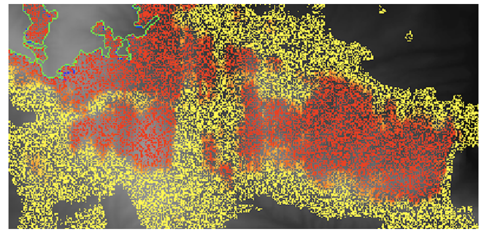

Figure 4: Hot spot southeast of Beaver Creek fire on July 16, 2016.

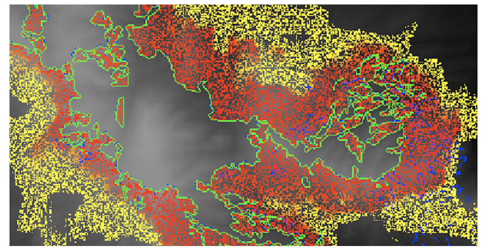

Figure 5: The same location as Figure 4, 14 days later.

Locations with Unique Characteristics

Figure 3 displays a subsection of FireCast's output containing examples of areas with dense clusters of pixels that were incorrectly predicted to burn. These clusters contribute to poor statistical results, although in a visual setting, we are able to make alternate conclusions about them.

Areas A and B are examples of clusters of pixels that are not within the next burn perimeter, but are completely surrounded by both the current and next day perimeters. We refer to these areas as *fire islands*. Area C is an example of what we refer to as a *fire valley*, a collection of pixels that are almost entirely surrounded by the fire perimeter, but are still open to land outside of the perimeter. Both fire islands and fire valleys are unique formations to be aware of as they contribute to worse statistical evaluation figures, though are areas that a fire strategist would likely not make an effort to save when developing a tactical suppression strategy.

Locations that Eventually Burned

When visually evaluating the testing fire, a unique trend appears in the output images over time. As the model progresses through the input perimeters, various dense clusters of pixels where $p \geq 0.5$ develop outside of the next day's perimeter, and are present in all outputs once they form. We refer to these clusters as *hot spots*. Hot spots are initially viewed as a gross over-prediction by the model, as they are not contained inside of the next day's perimeter.

Recall from §6.3, there is a significant data gap of 14 days between consecutively mapped fire perimeters for the testing fire. Figure 4 displays the southeast corner of the output image for the testing fire on the last day before the data gap. A large hot spot extends from the initial perimeter to the east, with an actual length extending over 9 kilometers along the northern ridge of a mountain. Figure 5 shows the output image for the exact same location on the first day after the data gap, which includes the green perimeter for August 2. Although this hot spot is initially evaluated as a large over-

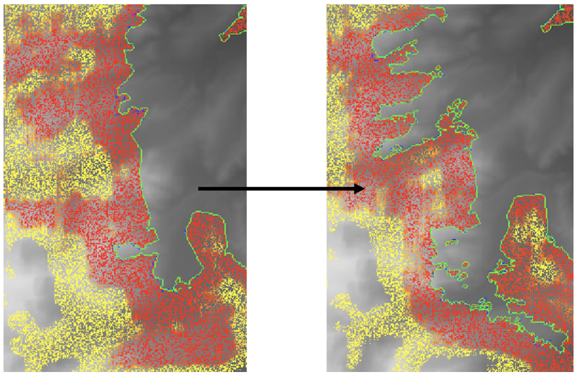

Figure 6: Western side of Beaver Creek fire. Image on left is July 16, 2016 and image on right is on August 2, 2016. This shows the hot spots to the west of the fire eventually burned.

prediction, the fire does eventually burn the entire area, validated by the perimeter two weeks later, despite the model not having access to atmospheric data for the data gap.

Figure 6 contains more examples of hot spots that formed around the perimeter before the data gap. Both images in Figure 6 are of the same location along the western edge of the fire perimeter, where the left is the last day before the data gap and the right is the first day after the data gap. We notice two main areas along the western side of the fire that contain significant hot spots, and those areas are also eventually overtaken by the fire, similar to the southeast corner. Although hot spots contribute to a weaker statistical analysis of the model, visual analysis unveils useful potential for fire strategists to identify high-risk areas earlier, possibly enabling earlier evacuations.

7 Discussion

FireCast is a novel algorithm applying supervised machine learning to effectively predict areas with a high-risk of burning as a wildfire progresses. AI and geospatial techniques are applied to a new application context in an attempt to help mitigate the costly and life threatening consequences of wildfires. Our model is designed to support flexible and informed decision-making by experienced fire fighters working to contain a fire. FireCast's predictive power gives visual insights to where the fire is likely to spread in the future instead of just displaying past fire data and providing weather forecasts.

Although Farsite is widely adopted by fire fighters and fire strategy experts, it is extremely computationally expensive and requires a large number of input variables unique to each different location. The process of initiating a Farsite simulation on any landscape requires a minimum input of five files including a land cover classification of fuels. FireCast only requires four different kinds of input data to simulate where a fire is likely to spread, all of which are in the public domain. Table 2 provides a direct comparison of the types of input data needed for each model.

Instead of performing predetermined calculations from manually collected input layers, machine learning within FireCast allows the system to learn important correlations. The reduced number of necessary and variable data, along with no need to generate both landscape and canopy cover files makes FireCast faster, less complex, and easier to implement than Farsite, while also outperforming Farsite with regards to total prediction accuracy, recall, and F-score. Visual evaluation shows FireCast's ability to make general predictions of high-risk areas for fire spread up to two weeks prior to the fire actually spreading to those areas.

FireCast	Farsite
Landsat8	Fuel Model
DEM	DEM
Fire Perimeters	Canopy Cover
Weather/Wind	Weather/Wind
	Adjustment File
	Fuel Moisture
	Conversion File*
	Fuel Model File*
	Fire Acceleration File*
	Fire Perimeters*

Table 2: FireCast vs. Farsite input variables. Optional input files (*).

8 Limitations, Future Work, and Conclusion

As aforementioned, FireCast is limited by the availability of appropriate training data. Our evaluation relied on methods such as weather interpolation and data augmentation to generate a sufficient amount of training data. We also ignored the affect that fire fighter activity could have on the training data since we do not have access to this information.

Future work for this project primarily involves gathering more perimeter and historical data for different regions worldwide, and exploring different resolutions of input variables. For this work, we manually gathered every fire perimeter in the public GeoMAC dataset for which there were consecutively mapped daily fire perimeters in the Rocky Mountain Region since 2013, when Landsat8 began to record images. While we assume the technique is applicable to different geographic regions, validation will require training and test data from regions other than the Rocky Mountain Region.

We present FireCast, a novel approach to predicting the spread of wildfire that applies supervised machine learning techniques to wildfire prediction. FireCast is lightweight and computationally inexpensive when compared to current popular models. Statistical analysis reveals FireCast learns patterns of wildfire spread from observing historical wildfires, and visual analysis of FireCast's output gives insight to the predictive power of the algorithm to identify high-risk areas of fire spread up to two weeks into the future. FireCast already outperforms current modeling software, and with more training data, is expected to increase in accuracy and scale to a variety of new regions with different physical features.

Acknowledgements

The authors would like to thank Colorado College undergraduate researcher Nick Crews, University of Waterloo Professors Kate Larson and Tim Brecht, and UC Berkeley Professor John Radke for their supervision, support, and contributions.

References

[Barker, 2008] Terry Barker. The economics of avoiding dangerous climate change. an editorial essay on the stern review. *Climatic Change*, 89(3):173–194, Aug 2008.

[Carlson and Ripley, 1997] Toby N. Carlson and David A. Ripley. On the relation between ndvi, fractional vegetation cover, and leaf area index. *Remote Sensing of Environment*, 62(3):241–252, 1997.

[Ding *et al.*, 2018] Peng Ding, Ye Zhang, Ping Jia, and Xuling Chang. A comparison: Different dcnn models for intelligent object detection in remote sensing images. *Neural Processing Letters*, Jun 2018.

[Finney *et al.*,] Mark Finney, Isaac C. Grenfell, and Charles W. McHugh. Modeling containment of large wildfires using generalized linear mixed-model analysis. *Forest Science*, 55(3):249–255.

[Finney, 1994] Mark A. Finney. Farsite: a fire area simulator for fire managers. *The Biswell symposium: fire issues and solutions in urban interface and wildland ecosystems*, pages 55–56, Feb 1994.

[Finney, 1998] Mark A. Finney. *FARSITE: Fire Area Simulator—Model Development and Evaluation*. US Forest Service, Ogden, Utah, 1998.

[Forghani *et al.*, 2007] Alan Forghani, Bob Cachet, John Radke, Mark Finney, and Bret Butler. Applying fire spread simulation over two study sites in california lessons learned and future plans. *IEEE Interantional Geosciences Remote Sensing Symposium*, pages 3008–30013, 2007.

[Houtman *et al.*, 2013] Rachel M. Houtman, Claire A. Montgomery, Aaron R. Gagnon, David E. Calkin, Thomas G. Dietterich, Sean McGregor, and Mark Crowley. Allowing a wildfire to burn: estimating the effect on future fire suppression costs. *International Journal of Wildland Fire*, 22:871–882, 2013.

[Kocsis and Szepesvári, 2006] Levente Kocsis and Csaba Szepesvári. Bandit based monte-carlo planning. In Johannes Fürnkranz, Tobias Scheffer, and Myra Spiliopoulou, editors, *Machine Learning: ECML 2006*, pages 282–293, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[Kourtz and O'Regan, 1971] Peter Kourtz and William G. O'Regan. A model for a small forest fire...to simulate burned and burning areas for use in a detection model. *Forest Science*, 17(2):163–169, 1971.

[Lauer *et al.*, 2017] Christopher J. Lauer, Claire A. Montgomery, and Thomas G. Dietterich. Spatial interactions and optimal forest management on a fire-threatened landscape. *Forest Policy and Economics*, 83:107–120, Oct 2017.

[Lin *et al.*, 2019] Zhonghie Lin, High H. T. Liu, and Mike Wotton. Kalman filter-based large-scale wildfire monitoring with a system of uavs. *IEEE Transactions on Industrial Electronics*, 66(1):606–615, Jan 2019.

[Maggiori *et al.*, 2017] Emmanuel Maggiori, Yuliya Tarabalka, Guillaume Charpiat, and Pierre Alliez. Convolutional neural networks for large-scale remote-sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(2):645–657, Feb 2017.

[O'Connor *et al.*, 2016] Christopher D. O'Connor, Matthew P. Thompson, and Francisco Rodríguez y Silva. Getting ahead of the wildfire problem: quantifying and mapping management challenges and opportunities. *Geosciences*, 6(3):35, 2016.

[Radke, 1995] John Radke. Modeling urban/wildland interface fire hazards within a geographic information system. *Geographic Information Sciences*, 1(1):9–21, 1995.

[Reidmiller *et al.*, 2018] David R. Reidmiller, Christopher W. Avery, David R. Easterling, Kenneth E. Kunkel, Kristin L. M. Lewis, T. K. Maycock, and Bradley C. Stewart (eds.). *Impacts, Risks, and adaptation in the united States: Fouth National Climate Assessment, Volume II*. U.S. Global Change Research Program, Washington, DC, USA, 2018.

[Simard *et al.*, 2003] Patrice Y. Simard, Dave Steinkraus, and John C. Platt. Best practices for convolutional neural networks applied to visual document analysis. *Proc. Int. Conf. Document Anal. Recognit.*, 3:958–962, Aug 2003.

[Stephens *et al.*, 2018] Scott L. Stephens, Brandon M. Collins, Christopher J. Fettig, Mark A. Finney, Chad M. Hoffman, Eric E. Knapp, Malcolm P. North, High Safford, and Rebecca B. Wayman. Drought, tree mortality, and wildfire in forests adapted to frequent fire. *BioScience*, 68(2):77–88, Feb 2018.

[Subramanian and Crowley, 2018] Sriram G. Subramanian and Mark Crowley. Combining mcts and a3c for prediction of spatially spreading processes in forest wildfire settings. *Canadian AI*, LNAI 10832:285–291, 2018.

[Séro-Guillaume *et al.*, 2008] Olivier Séro-Guillaume, Sepehr Ramezani, Jonathan Margerit, and Didier Calogine. On large scale forest fires propagation models. *Int. J. Thermal Science*, 47(6):680–694, 2008.

[Westerling *et al.*, 2006] Anthony L. Westerling, Hugo G. Hidalgo, Daniel R. Cayan, and Thomas W. Swetnam. Warming and earlier spring increase western u.s. forest wildfire activity. *Science*, 313(5789):940–943, Aug 2006.

[Zhang *et al.*, 2016] Qingjie Zhang, Jiaolong Xu, Liang Xu, and Haifeng Guo. Deep convolutional neural networks for forest fire detection. In *2016 International Forum on Management, Education and Information Technology Application*. Atlantis Press, 2016.