CSCI 1430 Final Project Report:
Wildfire Risk Assessment Using Satellite Imagery

Team name: Anna Dai, David Fryd, Sophie Ljung, Stella Ljung.
TA name: Byron Butaney. Brown University

Abstract

Wildfires pose a significant threat to ecosystems, human lives, and infrastructure. Early detection and accurate prediction of wildfire risks are crucial for effective firefighting and mitigation strategies and are valuable metrics used in analyzing the effects of the climate crisis. In recent years, deep learning techniques, particularly Convolutional Neural Networks (CNNs), have demonstrated remarkable success in various computer vision tasks. Our project focuses on leveraging CNNs to detect and assess wildfire risks based on satellite imagery. The proposed methodology involves the development of a CNN-based model capable of analyzing satellite images to identify regions at high risk of wildfires. The model is trained to incorporate information such as vegetation density and land cover types. Our dataset combines a pre-existing Canadian collection of satellite images, and we have augmented it by collecting satellite imagery across the US of locations that have experienced wildfires in 2023. By extracting meaningful geospatial features from the input data, the CNN learns to recognize patterns indicative of potential fire-prone areas.

1. Introduction

Climate change has driven a global increase in wildfires. 2020 was a record-setting year for wildfires in California, with almost 10,000 fires burning over 4 million acres of land. Wildfires present a substantial danger to ecosystems, human lives, and infrastructure. Our interest in wildfire risk prediction stems from an investment in combating the climate crisis: we hope to critically consider AI tools that enable efforts in preventing, mitigating, and managing wildfires more effectively.

The impact of wildfire risk assessment tools like ours is significant. Human intervention continues to cause significant changes to landscapes around the world. Wildfire predictions based on historical data fail to consider the ways that areas might rapidly change due to deforestation, urban development, and climate change. The ability to infer the risk of wildfires based on the latest landscape conditions and alterations is critical in protecting ecosystems, wildlife, and human lives from the consequences of climate change.

Wildfire prediction is complex. They occur in vast areas and exhibit complex spatial and temporal dynamics. Predicting the behavior and spread of wildfires requires capturing and analyzing fine-grained details about the landscape, including wind patterns, fuel moisture content, and terrain characteristics. In addition, wildfires can range from small localized incidents to large, rapidly-spreading disasters. Integrating and analyzing diverse data sources with varying formats and quality presents a significant challenge.

We focused on performing risk assessment purely through landscape analysis. While more complete risk-assessment algorithms incorporate many non-image-based factors, we sought to explore and maximize the accuracy of purely image-based prediction systems. Our model and process would fit very well serving as the landscape-analysis component in a future fully-featured wildfire risk assessment tool.

2. Related Work

Kaggle We used a pre-labelled existing dataset of about 44 thousand satellite images from Canada to train our wildfire risk model. Roughly 50% of images were sampled from locations where a wildfire has occurred and the other 50% sampled from locations with no history of wildfire. [1]

National Interagency Fire Center The National Interagency Fire Center stores geographical data on wildfires from 2014 to 2023. We used the 2023 data, filtering on fire-reported incidents, to augment the Kaggle dataset with imagery from the US. [2]

Google Earth Engine Google Earth Engine is a geospatial processing service that allows users to query for satellite imagery given latitude and longitude. We used its API to augment our model's training and testing data given coordinates of prior wildfire incidents in 2023. [3]

Mapbox Mapbox API is another geospatial processing service that also allows users to query for satellite imagery given latitude and longitude. We used its API to download more than 50 thousand square kilometers of satellite imagery of Oregon to make wildfire risk predictions across the state (images fed into model). We switched from Google Earth Engine because of issues with corrupted satellite data initially experienced when working with GEE. We also used Mapbox to generate the heatmap visualizations of our results. [4]

United States Forest Service We referenced the 2018 visualization of "Burn Probability" generated by the United States Forest Service as a point of comparison for our own wildfire risk heat map. [5]

3. Method

Our project was completed in 4 general steps. We first augmented an existing dataset of Canadian satellite imagery with our own American satellite images of places that experienced wildfires in 2023. We then built a convolutional neural network that evaluates wildfire risk given an RGB satellite image of dimensions 224×224, and trained it on the augmented dataset. Next, we downloaded images covering the state of Oregon to be later used for wildfire risk assessment. Finally, we ran the Oregon-based images through our model and constructed a heat map visualization representing our assessment of wildfire risk in the state of Oregon.

3.1. Data Augmentation

The pre-labelled dataset is comprised exclusively of satellite imagery from Canada. To make our model more robust to geographical variety and non-Canadian landscapes, we augmented the dataset with satellite imagery from the US to introduce diversity in vegetation and development. We first consulted the Natinoal Interagency Fire Center for specific latitudes and longitudes of 2023 American wildfires, and used Google Earth Engine to download images of these locations. We then split the downloaded data among test and train data for the model evenly. Collecting images of locations with the absence of a history of wildfire is far more complex, and is outside of the scope of the augmentation we performed. Figure 1 contains a visualization of the data augmentation pipeline.

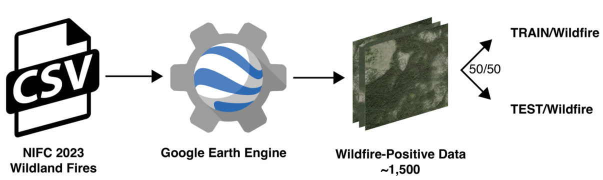

Figure 1. Data Augmentation Pipeline

3.2. Model Architecture

Our CNN Binary Image Classifier consists of 5 convolution blocks, a Flatten layer, and 2 Dense layers. Our model performs Stochastic Gradient Descent with momentum on 917,655 total trainable parameters. We chose SGD as our optimizer for its efficiency, scalability, and convergence rate. SGD is computationally efficient, making it well-suited for a large dataset. SGD is scalable, because it can handle a large number of input features, including images with many pixels. Ultimately, SGD converges to a good solution in a short amount of time, reducing the computational resources required to train our model. Using general observation to select our training hyperparameters, we trained for 15 epochs with a learning rate of 0.001 and momentum of 0.01. See figure 2 for a visualization of our model's architecture.

Each convolution block contains a 2D Convolution layer with Leaky Relu activation directly followed by a Max-Pooling layer. We chose to use an ascending number of filters in each convolution layer: the first 2D convolution layer employs 16 filters, the second employs 32 filters, and the following three layers employ 64 filters.Increasing the number of filters increases the number of abstractions our network is able to extract from image data. Our images capture the color, density, and distribution of both vegetation and urban infrastructure. The raw data from any given image may contain noise that we do not want our model to learn as a classifying feature. Our first convolution layer has the least number of filters so that we can extract the most relevant, primitive features from the raw pixel data directly from an input image. Once the first convolution layer extracts these features from the input image, our network uses an ascending number of filters to increase the depth of the feature space and learn more complex interactions between global abstract features. Shrinking the feature space before feeding it to the dense layers of our network helps the dense layers more easily learn to map the reduced feature space to the output labels, improving model performance through better classification accuracy and diminishing risk of overfitting.

Each 2D convolution layer employs a Leaky ReLU activation function. This nonlinear activation function helps address the vanishing gradient problem that can occur with traditional ReLU activation. With the introduction of a small non-zero slope, gradients may still flow through the network for negative input values, improving our model's ability to learn and generalize complex features.

We follow the 2D convolution layer in each block with a Max Pooling layer using a stride of size 2. This reduces the dimensions of the input feature map from the previous layer by one fourth. Downsampling the feature map decreases the number of parameters, decreases likelihood for overfitting, and increases our model's computational efficiency. Max pooling enables better learning of high-level features, such as trees or mountains in our images. In addition, max pooling provides local translation invariance, meaning the model can learn features regardless of their location in the image. This proves especially useful for our images of natural scenes,

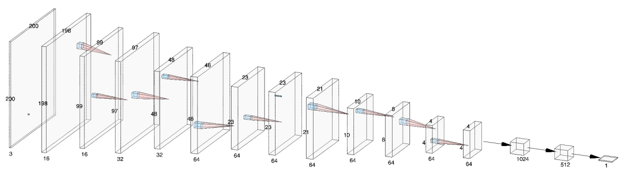

Figure 2. Model architecture

since patterns in vegetation appear in different positions and orientations.

We chose to include two dropout layers with a 20% dropout rate following the third and fifth convolution block. The dropout layers choose random neurons to "drop out" of the network by setting their weights to zero, keeping the model from relying heavily on any specific neuron, or feature, in the network. Regularization through dropout increases model performance by making it more robust to new, unseen data.

After performing our convolution, we flatten the 2D output into a 1D vector, enabling our fully-connected dense layers to classify the extracted features. We use two dense layers in our final stages of training. Each dense layer performs matrix multiplication in order to preserve the weights from the previous neurons while decreasing the dimension of the vector. Our first dense layer has a size of 512, which decreases the dimension of our 1D feature vector to 512. Since our classifier is binary, we use a sigmoid activation function that will output one value in the range of 0 to 1 in our final dense layer.

3.3. Oregon Data Collection

In order to assess the risk of wildfire in the state of Oregon, we needed to collect satellite image data spanning the entire landmass of the state. The general approach was to collect a list of points (longitude and latitude), then download images of those points to be later fed into the model.

3.3.1 Sampling Points

In an ideal world, we would have sampled points every "X" kilometers as we traversed the area of the state and downloaded images of appropriate dimensions so that we would have 100% image coverage of the state. In reality, we were limited by the download cap of satellite images from our chosen API, where we were restricted to downloading a total of 50,000 images before being charged. Therefore, we needed to generate an even and representative distribution of points in order to achieve the best assessment of wildfire risk across the entire state.

We generated points using Poisson-disc sampling, allowing us to define a minimum distance between points and fine-tune our data set for an even distribution. However, a challenge we faced while applying this method was with the geographic scale with latitude. Due to the Earth's spherical shape, the same difference in longitude represents different actual distances depending on the latitude of a given point. This variation requires us to convert our minimum distance from meters to degrees for Poisson-disc sampling using the Haversine formula, factoring in the latitude of the points we sampled. Figure 3 shows the results of a single iteration of sampling.

We tried different distributions and disc-radii to keep within the limitations of our download cap. Our final highest-density result consisted of 5,518 points distributed across Oregon. We later discuss the differences in sampling density (Look ahead to Figure 7).

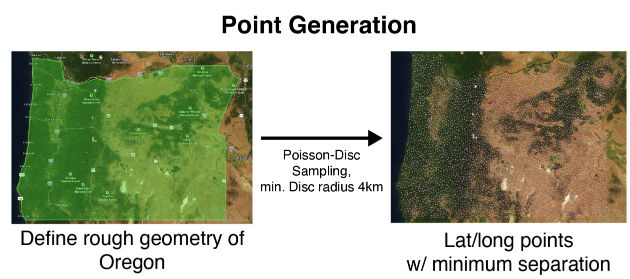

Figure 3. Generate points using Poisson-disc sampling

3.3.2 Downloading Image Clusters

After generating a set of points distributed across the state of Oregon, we iterated over each of these points in order to download images of these locations. Instead of just downloading a single image per point, we download 9 1x1km images arranged in a square matrix centered around the point. This gives us a more accurate representation of the surrounding area making our predictions more robust to noise and false-positives/negatives. Each set of images for a given point is associated with the sampled latitude/longitude, and the predictions of the model across all 9 of those images contribute to the wildfire risk score of the given point. Figure 4 shows the cluster download process. For the densest sampling we did with 5,518 points, we downloaded a total of 49,662 images.

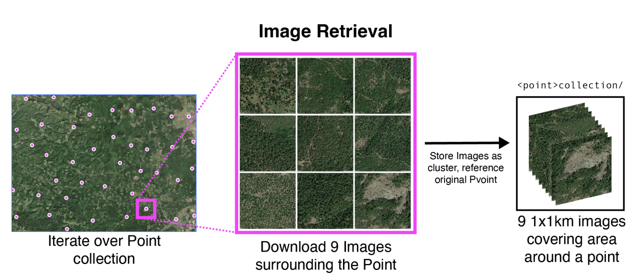

Figure 4. For each point, 9x9 set of images is downloaded, centered around the point and grouped together by the central point.

3.4. Risk Assessment Visualization

After collecting all of the image clusters for all of the points, we ran each cluster of images through our CNN. The number of "wildfire-positive" votes out of 9 was calculated for each point, resulting in an output .geojson file predicting wildfire risk across the entirety of the sampled space. Using the .geojson output file, we constructed a heatmap where each point contributes an "intensity" to the heatmap proportional to the number of wildfire-positive votes it received from the model.

Figure 5 shows how each of the point image clusters contributed to the final visualization of wildfire risk.

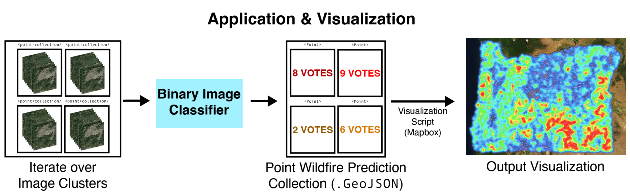

Figure 5. Pipeline: Point Image Cluster to Visualization

4. Results

We can analyze our results in two ways: Quantitatively by evaluating of our underlying CNN, and qualitatively comparing our predictions with existing predictions made by external sources.

4.1. Model Performance

Our model was trained using a learning rate of 0.01 and a momentum of 0.0001. The learning rate determines the step size at which the model updates its parameters during training, while momentum helps accelerate the optimization process by adding a fraction of the previous parameter update. These values were chosen to strike a balance between learning efficiency and stability during training.

Our model achieved 96.4% accuracy over 15 epochs. Table 1 contains more granular results obtained from our model's performance evaluation. It provides a comprehensive overview of the model's accuracy and loss values across 15 epochs. We exclude above epoch 15 because we did not see an improvement in performance after the 15th epoch, and exclude intermediate epochs for brevity.

To evaluate the performance of our model, we utilized two metrics: binary accuracy and loss as measured by binary cross-entropy. Binary accuracy measures the proportion of correctly classified samples, providing an indication of the model's overall prediction accuracy. On the other hand, binary cross-entropy loss quantifies the dissimilarity between predicted probabilities and true labels, allowing us to assess the model's training progress and generalization capabilities.

Epoch	1	2	3	10	15
Accuracy	0.904	0.922	0.943	0.955	0.964
Loss	0.259	0.202	0.157	0.128	0.105

Table 1. Results: accuracy and loss measured across 15 epochs. Both accuracy and loss continuously improve with time.

4.2. Oregon Wildfire Analysis

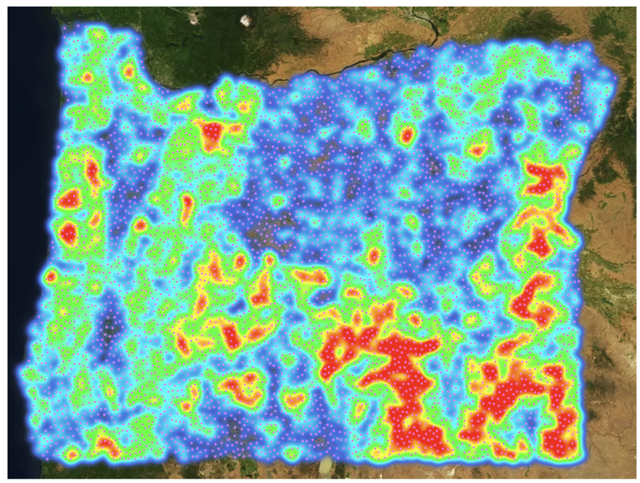

Figure 6. Our assessment of wildfire risk in the state of Oregon. Pink dots represent the 5,518 sampled locations. Redder color indicates areas with a higher degree of wildfire risk. Each point contributes an "intensity" to the heatmap proportional to the number of wildfire-positive votes it received from the model.

Figure 6 represents the cumulative wildfire risk assessment we produced for the entire state of Oregon. Figure 7 shows our predictions in comparison to a Burn Probability Map created in 2018 by the United States Forest Service.

Our results were somewhat similar to predictions made by the U.S. Forest Service, with a few notable discrepancies. Figure 7 shows a visual representation of areas of high/low correspondence between our predictions. In essence, we concur with most areas deemed "high risk" by the USFS, but disagree in some areas that are labeled "low risk" by the USFS (mainly the coast of Oregon). We discuss our theories for these discrepancies in the next section.

4.3. Technical Discussion

It is important to note that the United States Forest Service used more than just landscape analysis to generate their predictions. Historical and meteorological analysis play a large part in fully-fledged wildfire prediction algorithms. We expect that the discrepancies in the visualization, for example the high-risk we predict in the coastal region, stem from the lack of consideration for non-image based data (i.e. extremely high average annual precipitation in the coastal

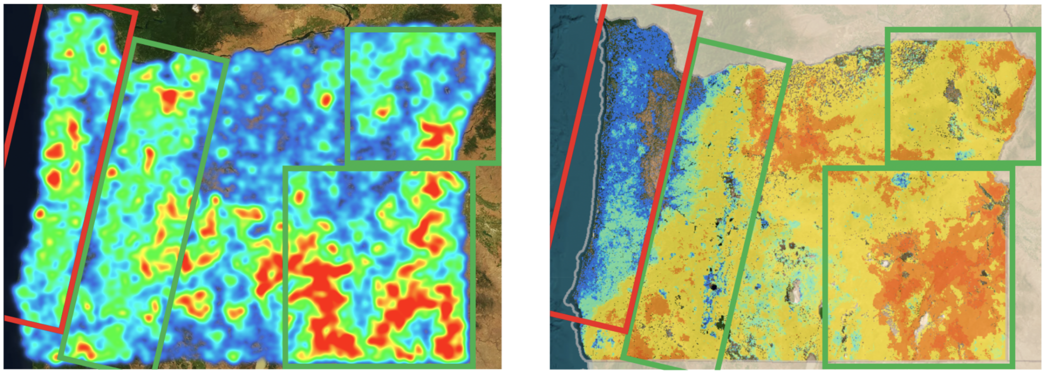

Figure 7. *Left:* Our assessment of wildfire risk in the state of Oregon. *Right:* A 2018 visualization of "Burn Probability" generated by the United States Forest Service.

region). Figure 9 demonstrates this perfectly- if you were to use the precipitation map of Oregon to alter our predictions (dampening areas with high precipitation), our outputs are *strikingly* similar to predictions made by the USFS. Considering the fact that we excluded all other non-image-based factors, we were pleased to generate impressively similar results based on image data alone.

As mentioned previously, wildfire risk assessment is highly complex, and fully-fledged algorithms account for many non-image based factors including weather, vegetation and human activity. Because our model is trained only on satellite imagery, it cannot consider non-image based factors when predicting wildfire risk. Given more time, we would have liked to incorporate weather statistics (mainly temperature, humidity, and wind speed) acquired via Python weather APIs into a larger deep learning model to supplement our satellite imagery computer vision model.

Additionally, we chose to limit our wildfire analysis to North America, where terrain is categorically comparable. We chose to analyze Oregon specifically due to its history of wildfires and smaller area compared to other wildfire-prone states, such that fewer samples would yield a more representative result. Its geometry is close to rectangular, which simplified the process of defining borders for our sampling purposes.

The area of the state of Oregon is roughly 250,000 square kilometers in area. In our first iteration, we sampled 734 (Poisson disc radius 11 kilometers) points across the entire state, which produced a dataset of 6,606 images at 2.5% coverage. The heat map visualization generated by this sparse point set yielded a spotty visualization. This made it difficult to recognize patterns and draw conclusions because of the large amount of interpolation in area between the sampled points. The apparence of this difference is visible in Figure 8. We then decreased the disc radius to four kilometers to generate a total of 5,518 points, which yielded 49,662 images at about 20% landmass coverage.

With this modification, we saw areas of high risk emerge in the resultant heat map. The increase in data produced a much more legible risk map that had comparable results to the risk map generated by the United States Forest Service (see 8). However, our image queries were limited by the number of requests we could make to the image download API (which was capped at 50,000 free requests per account).

4.4. Socially-responsible Computing Discussion via Proposal Swap

Our group gathered data from wildfire satellite imagery in the United States. We then augmented the Canadian dataset in order to include more pictures of predicted American wildfires. Next, we trained and tested the model on the augmented dataset so the model was more robust to geographical changes in the environment. In addition, we also augmented the data in our preprocessing through shifting the images, flipping the orientation, and changing the scale. The augmentation of the dataset in our model attempted to counteract the dependency the neural network may have on just Canadian geography. Unfortunately we did run into issues that the Oregon dataset had different geographical properties, especially with more mountainous regions, than the augmented dataset we trained on. In the future, we would work to add more diverse geographical images in our dataset so that the model is robust to all different geographical areas.

Adding in extra data about the population density and fire safety access would make the model more robust to real-life predictions of wildfires. The proximity of a fire department shows how quickly a wildfire could be extinguished, therefore decreasing the likelihood of a massive wildfire. In addition, the population density correlates to the risk of damage within that region. With a higher population density the wildfire will have a greater direct impact on people's lives, causing the wildfire to be much more dangerous. Overall, it is true that both these properties do not directly impact the risk of a wildfire within a certain region. Instead, they give additional context to the harm a wildfire may possess if it's nearby a large town (impacting lives) and/or far from

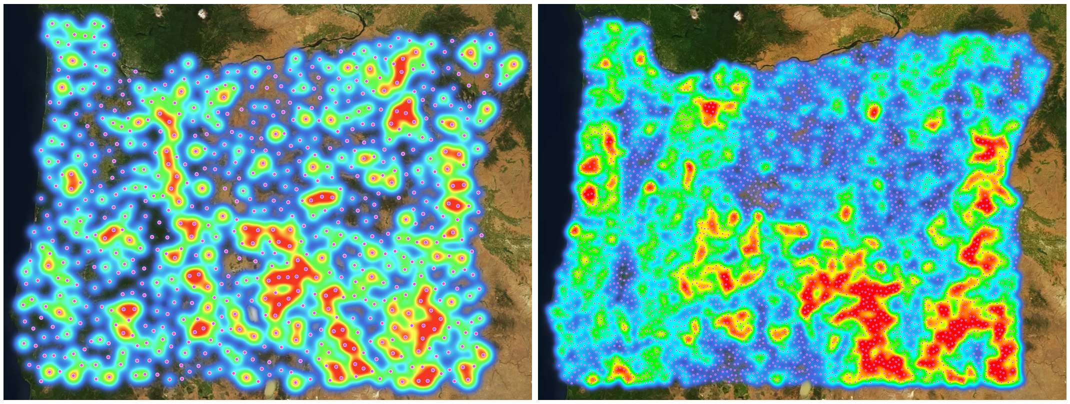

Figure 8. Difference between sampling strategies. *Left:* **734** points, 6,606 total images. *Right:* **5,518** points, 49,662 total images.

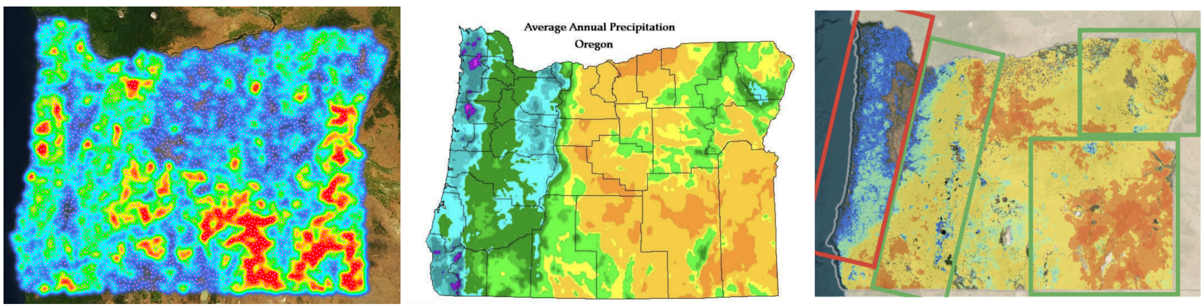

Figure 9. *Left:* Our assessment of wildfire risk. *Center:* Annual Avg. Precipitation in Oregon (bluer is higher precipitation) *Right:* Areas of close-match (green) and discrepancies (red). If you overlay the precipitation of Oregon onto the predictions we made, the predictions are *far more* similar, emphasizing the importance of non-image based factors in more complete prediction algorithms.

fire safety access (can spread more easily). This information will be used less for predicting the likelihood of a wildfire, but instead more towards interpreting the results a wildfire would have on that specific region.

The augmented dataset provides greater context to prediction of wildfires in North America. Since the model is trained and tested specifically on the combined dataset of Canadian and United States predicted wildfires, it has the highest accuracy in detecting predicted wildfire in these regions. With a greater bandwidth of resources in this project, we would have liked to further augment the dataset to include other regions as well. Time-permitting, we would have included other regions across the globe so that this model could be generalized to predict wildfires worldwide. We decided to use the augmented dataset in Canada and the United States in order to ensure our model was working on a small scale before expanding it to generalize it to other regions. Our American dataset was found by taking images using Google Satellite Imagery of locations of previous wildfires. In this dataset, we did not focus on American areas with no-wildfire. Thus, we agree there is some bias within our model to detect Canadian wildfire/no-wildfire regions the best, followed closely by the American regions, and then finally globally. In the future, we would work to include a larger dataset with greater global data augmentation so that the model can be applied more broadly.

5. Conclusion

Wildfires present a notable danger to ecosystems, human lives, and infrastructure. Detecting and predicting wildfire risks at an early stage is vital for effective firefighting and mitigation strategies. In recent times, deep learning methods, specifically convolutional neural networks, have shown remarkable achievements in diverse computer vision applications. Our project aimed to utilize CNNs to identify and evaluate wildfire risks. Our assessment of wildfire risk based on our model's predictions was comparable to the 2018 visualization of burn probability generated by the United States Forest Service, even though our model only accounts for geospatial features.

Such a model for wildfire risk detection can greatly assist firefighting efforts, especially in the Western US, by providing valuable insights and actionable information. For example, a model trained to detect wildfire risks can analyze high-resolution satellite images in near real-time, identifying areas with a high likelihood of wildfires. By providing early detection, the model enables firefighting agencies to respond promptly and allocate resources effectively, increasing the chances of containment before fires become uncontrollable.

Furthermore, our model can help prioritize resource allocation. Firefighting agencies could focus their efforts on regions flagged as high-risk by the model, ensuring that limited resources such as firefighters, equipment, and aircraft are deployed strategically to the areas most in need. Beyond image data, a full fledged wildfire risk model could also incorporate data such as building records and infrastructure details, to assess the vulnerability of structures within wildfire-prone areas. By identifying buildings or critical infrastructure at higher risk, firefighters can prioritize protection efforts and allocate resources accordingly, potentially reducing property damage and saving lives.

By leveraging the capabilities of a CNN-based wildfire risk detection system, firefighting efforts in North America can become more proactive, efficient, and targeted. This technology provides crucial insights into identifying at-risk areas, prioritizing resources, monitoring fire behavior, and protecting communities and infrastructure. Ultimately, it can help mitigate the devastating impacts of wildfires, safeguard lives, and minimize property loss.

References

[1] Abdelghani Aaba. Wildfire Prediction Dataset (Satellite Images) - Satellite images of areas that previously experienced wildfires in Canada, howpublished= https://www.kaggle.com/datasets/abdelghaniaaba/wildfire-prediction-dataset. 1

[2] National Interagency Fire Center. 2023 Wildland Fire Incident Locations to Date - Point locations for all reported wildland fires in the United States in 2023 to date., howpublished= https://data-nifc.opendata.arcgis.com/datasets/nifc::2023-wildland-fire-incident-locations-to-date. 1

[3] Google. Earth Engine, howpublished= https://earthengine.google.com/. 1

[4] Bonnie Bogle Eric Gundersen. Mapbox API, howpublished= https://api.mapbox.com. 2

[5] United States Forest Service. United states forest service burn probability map via oregon explorer. https://tools.oregonexplorer.info/oe_htmlviewer/index.html?viewer=wildfire. 2

Appendix

Team contributions

Anna and David are captioning. Capstone contributions revolve around the wildfire risk analysis of Oregon as well as its visualization.

Anna (Capstoning) Helped David construct North American wildfire dataset, and worked with them to visualize our model's wildfire risk analysis of Oregon. Wrote the code to process Oregon satellite imagery, pass the images through the trained model, and execute a confidence vote at each cluster.

David (Capstoning) Wrote the code for querying Google Earth Engine based on NIFC .csv dataset, and storing the data on Google Drive for augmentation. Wrote the code for geospatial Poisson-disc sampling for generating point samples in Oregon. Wrote the code to calculate matrix cluster points, and interface w/ Mapbox to download clusters of images at each pt. Wrote web-app visualization tool using Mapbox. Made Poster and Diagrams for Poster.

Sophie Worked with Person 4 to design CNN binary image classifier. Fine-tuned hyperparamaters to train model for optimal performance.

Stella Worked with Person 3 to design CNN binary image classifier. Implemented data pre-processing and run script.