
Predicting the Occurrence of Wildfire using Convolution Neural Network

Aashi Goyal

Department of Computer Science
University of Southern California
Los Angeles, CA, 90007
aashiary@usc.edu

Astha Modi

Department of Computer Science
University of Southern California
Los Angeles, CA, 90007
anmodi@usc.edu

Manali Maniyar

Department of Computer Science
University of Southern California
Los Angeles, CA, 90007
mmaniyar@usc.edu

Shalin Bhavsar

Department of Computer Science
University of Southern California
Los Angeles, CA, 90007
shalinbh@usc.edu

Abstract

In response to the need for enhanced wildfire prediction, this research focuses on harnessing the power of Convolutional Neural Networks (CNNs) to create standardized, compatible labeled satellite image datasets for various regions, starting with California and Canada. By harmonizing these datasets in terms of image properties and labeling, we have tried to ensure that models are transferable across regions, allowing training on one region's dataset and application to others. These models aim to learn the terrain patterns associated with wildfire occurrences on a global scale. Using CNNs, the study predicts the probability of wildfire risk from satellite imagery, aspiring to build generalized models applicable worldwide, transcending regional limitations. As new regions are incorporated, the datasets are incrementally standardized. This innovative framework aims to identify high-risk areas globally, enabling more effective wildfire prevention. The paper underscores the role of deep learning and CNNs in timely wildfire detection and highlights the significance of fine-tuning model parameters for improved prediction accuracy.

1 Introduction

1.1 Motivation

The motivation behind this research is grounded in the growing significance of wildfire risk assessment and mitigation within the realms of environmental conservation and human safety. Wildfires represent a recurring and destructive natural disaster, and the ability to understand their patterns and forecast their occurrences is paramount for proactive measures. These devastating events have far-reaching impacts on ecosystems and human lives, underscoring the necessity to develop advanced predictive models that can assess wildfire risk with enhanced precision and reliability. In this pursuit, we aim to harness readily available geospatial data and state-of-the-art machine learning techniques to construct a model capable of predicting wildfire risk based on satellite imagery. Such a model has the potential to facilitate the identification of high-risk areas, enabling timely preventative measures and ultimately contributing to the preservation of natural resources and the protection of communities. This research aligns with Sustainable Development Goal 13, "Climate Action," which emphasizes the urgent need to combat climate change and

its impacts, including natural disasters like wildfires. Additionally, it supports Sustainable Development Goal 15, "Life on Land," by aiming to protect, restore, and promote sustainable use of terrestrial ecosystems and halt biodiversity loss, which are often the direct victims of uncontrolled wildfires.

1.2 Problem Statement

Wildfires, as relentless and destructive forces of nature, necessitate our vigilant attention to safeguard lives and the environment. The ability to predict these fires is pivotal in preventing their catastrophic consequences. Traditional methods employed for monitoring and detecting forest fires have evolved over time with advancements in technology. The utilization of lookout towers and patrols was a common practice, but these approaches had inherent limitations, particularly in remote areas where early fire detection was challenging. The advent of satellites equipped with specialized sensors marked a significant turning point. These satellites had the capability to detect distinctive indicators of fires, such as heat and smoke, from high above the Earth, enabling faster and more accurate fire detection. Recent technological advancements have further refined our capacity to detect forest fires, incorporating innovations like drones equipped with sensors and ground-based monitoring systems. However, our research delves into an especially promising domain: deep learning, with a particular focus on Convolutional Neural Networks (CNNs), to further enhance the intelligence of our fire detection systems.

1.3 High-Level Outline of Goals and Approaches

Our principal objective in this project is to utilize CNNs to analyze images and determine the presence of fires. This analytical capability allows for early fire detection, which is of utmost importance for the protection of individuals, property, and the environment. However, the applications extend beyond early warning. Our goals encompass saving lives, minimizing damage, and preserving our invaluable natural resources.

1.3.1 Image-Based Approach

This approach pivots on the development of machine learning models that are trained to detect potential wildfire risks by analyzing satellite imagery and correlating it with historical data on wildfire occurrences, including precise latitude and longitude coordinates. By integrating real-time satellite data with past wildfire geolocations, the model aims to identify patterns and indicators of wildfire susceptibility in various regions. Such a preemptive model can be invaluable for mobilizing firefighting efforts and initiating evacuation plans, thereby minimizing the spread of fires and reducing their impact on ecosystems and communities. This data-driven, proactive approach enhances the precision of wildfire risk assessments and supports the efficient allocation of resources for fire prevention and control.

2 Related Work

In the domain of fire detection from satellite imagery, recent advancements in deep learning and image analysis have paved the way for more accurate and efficient methods. Several studies have leveraged innovative techniques to tackle the critical task of identifying active fires from satellite images. In this literature review, we delve into the few papers who have contributed significantly to the field by introducing novel approaches to fire detection.

Jing et al. [1] introduce an advanced scene image classification technique that merges the Alex-Net model with support vector machines in a deep convolutional neural network setting. It primarily focuses on harnessing features from Alex-Net's final layer to boost classification precision. Tested on ImageNet2012 and NUS-WIDE datasets, this approach showed notable improvement in classification outcomes. Future research directions include examining inter-category relationships to further refine scene classification.

Weitao Li et al. [2] present a novel approach to texture feature extraction using Gabor filters. Their method focuses on the selection of specific Gabor filters to create a more compact and efficient filter bank. This approach significantly reduces computational complexity and enhances classification performance in texture feature extraction. The study validates the effectiveness of this method through experiments on benchmark datasets and a real-world application, demonstrating improved classification accuracy and reduced filter count compared to traditional methods. This innovation in Gabor filter selection contributes to more effective and efficient texture analysis in various applications.

Priya et al. [3] use the Inception-v3 architecture to classify satellite images and detect the presence of fire.

The dataset contained satellite images as well as other images with textures similar to fire. The training data (481 images) consisted of satellite images containing active forest fires, with training pixels obtained from representative polygons containing fire and non-fire regions. The network was trained to distinguish between fire and non-fire using a standard CNN method. Local Binary Pattern (LBP) was applied to the corresponding image to detect fire-affected regions. 53 images were used for testing. The proposed method achieved a high accuracy for fire detection, with a weighted average of 98% accuracy based on the datasets used.

Spiller et al. [4] present deep learning techniques to detect spectral dependencies in wildfire situations using both spectral and spatial analyses. It utilizes models trained on data from Australia and Sicily and performs generalization tests in Oregon to assess the methodology's capacity to generalize across different ecosystems. The first two models (FC and 1D-CNN) are designed for spectral analysis and are noted for their spectral analysis capabilities, while the second two models (2D and 3D-CNN) are tailored for combined spatial and spectral analysis. The fire class was categorized with a perfect F1 score of 100% according to the FC and 1D models. Cross-validation on the training dataset produced mean precision, recall, and F1 scores ranging between approximately 0.97 and 0.98. The fully connected model demonstrated superior generalization capabilities, while the 3D CNN offered more refined classifications. Despite the high accuracy, the models exhibited a tendency to misclassify other classes and were prone to generating false alarms.

Zhang et al. [5] presents a spatial prediction model for forest fire susceptibility using a Convolutional Neural Network (CNN). The study used a dataset of 14 influencing factors, including meteorological, topographical, and vegetation factors, to map forest fire locations in Yunnan Province. The dataset was preprocessed and class imbalance was eliminated in the oversampling process. The CNN model was constructed using the Keras DL framework that uses TensorFlow as a backend. The model was trained and validated using the dataset, and the performance was evaluated using the test dataset. The data used in the study was obtained from the Climate Forecast System Reanalysis (CFSR) and the Yunnan Forest Fire Prevention and Control Headquarters. The study achieved an accuracy of 0.89 for forest fire susceptibility prediction using the CNN model.

The methodology used by Santopaolet al. [6] is based on a deep convolutional neural network to predict the risk level of wildfire from satellite data. The authors developed a custom Convolutional Neural Network for the purpose of pixel-wide short-term fire risk prediction. The system is based entirely on remotely sensed data, gathered from public satellite sources. The authors used a dataset consisting of satellite images of the surveilled area, along with the corresponding fire risk level labels. They achieved a MSE of 0.0002. CNN's prediction precision is demonstrated through case studies in Sicily and California.

3 Background on Deep Learning and Satellite Imaging

3.1 Deep Learning and Convolutional Neural Networks

Deep Learning, a remarkable field within artificial intelligence, has garnered immense attention in the research community due to its exceptional ability to tackle complex problems and generalize from data. At its core, Deep Learning employs a class of functional approximators known as Artificial Neural Networks (NN). These networks consist of layers, each containing numerous interconnected neurons. The process of training a neural network involves adjusting its internal parameters, known as weights, through a procedure called backpropagation. This training enables the network to approximate complex functions that link input data to output data, making it a powerful tool in solving various tasks.

Within the realm of Neural Networks, Convolutional Neural Networks (CNNs) stand out as a specialized class. CNNs are designed to address specific challenges in image analysis and recognition. They offer a unique feature: each neuron in a CNN has limited connectivity with the neurons in the previous layer. This local connectivity allows CNNs to create much deeper models while managing to keep the number of trainable parameters in check. The distinctive characteristic of local connectivity in CNNs is particularly advantageous for identifying spatial patterns in data. For instance, it can detect faces in photographs or objects within a background, making CNNs an ideal choice for tasks like image analytics and, notably, predicting fire risks.

The basic architecture of a CNN can be described as input layer which receives the input image, which is usually a 2D array of pixel values. Then is the convolutional layer which applies a set of learnable filters to the input image, which helps to extract features from the image. Each filter produces a feature map, which

is a 2D array of activation values. The third is the activation layer which applies a non-linear activation function to the output of the convolutional layer, which helps to introduce non-linearity into the model. The pooling layer downsamples the output of the activation layer, which helps to reduce the spatial dimensions of the feature maps and make the model more computationally efficient. The fully connected layer connects every neuron in the previous layer to every neuron in the current layer, which helps to learn complex patterns in the data. Lastly, the output layer produces the final output of the model, which is usually a probability distribution over the possible classes [5].

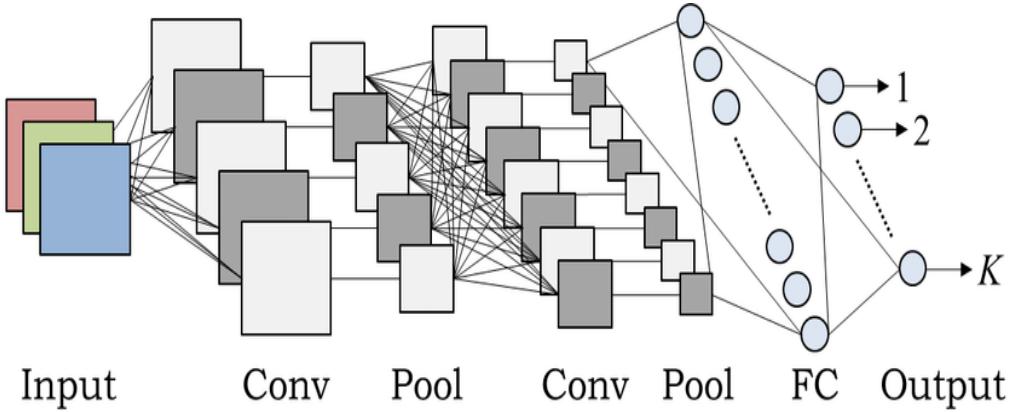


Figure 1: An example of CNN architecture

In simpler terms, Deep Learning harnesses the power of neural networks to understand and learn complex relationships in data, while Convolutional Neural Networks, a specialized type of neural network, excel at deciphering patterns in images, which is incredibly valuable when it comes to tasks like predicting the occurrence of wildfires based on visual data.

3.2 Satellite Imaging

Satellite Imaging is the process of gathering and evaluating data from satellites in Earth's orbit. These satellites are outfitted with cameras and sensors that record pictures and information about the Earth's atmosphere and surface. A valuable tool for many tasks, satellite imagery can be used to analyze land use, follow environmental changes, and monitor weather patterns, among other things. We have used these satellite images to generate a dataset with wildfire and no wildfire classes.

4 Data

The project involves creating standardized and labeled satellite image datasets for California to predict wildfire risk. These datasets are designed to match image properties and labeling of the Canada dataset used, enabling the development of generalized models for wildfire risk prediction. We are using the open dataset "Wildfire Prediction Dataset" for Canada and to maintain data consistency and avoid overfitting of the model over Canada's terrain, we have created a California wildfire dataset using the same method. By creating datasets that match in terms of image properties, labeling, etc. we ensure models are transportable across regions. Models can be trained on one region's dataset and applied to another region, since the data is compatible. This allows building generalized models to predict wildfire risk from satellite imagery, not restricted to just California or Canada. The models will learn terrain patterns that are associated with wildfire occurrence across regions. The data includes geospatial information, and the goal is to identify high-risk areas for wildfire prevention. To achieve this, a variety of machine learning and data analysis methods can be employed, such as convolutional neural networks (CNNs) for image classification, spatial analysis for terrain pattern recognition, and probabilistic models for risk assessment. The specifics of the data, metadata, and data extraction methods will provide a foundation for selecting appropriate techniques to enhance accuracy and predictive power.

4.1 Data Source

4.1.1 Canada Dataset

The open dataset used in this study originates from Canada's Forest Fires data [6], which is openly accessible through the Open Government Portal. The dataset comprises a collection of satellite images, each standardized to 350x350-pixel resolution. These images fall into one of two distinct classes: "Wildfire," encompassing 22,710 images, and "No wildfire", which includes 20,140 images. Here, wildfire indicates images of places where "Wildfire" has occurred in past and "No wildfire" means satellite image of places where wildfire has never occurred. However, the timeframe of this dataset is from 1972 - 2021. The dataset has been thoughtfully partitioned into distinct subsets for the purpose of training, testing, and validation, distributed as follows: the training set encompasses approximately 70% of the complete dataset, while the testing and validation sets each represent about 15%. The compilation of this dataset followed a systematic methodology. It involved the extraction of satellite images corresponding to specific geographical coordinates defined by longitude and latitude. These coordinates were associated with regions where wildfires, characterized by a burning area greater than 0.01 acres, were identified. To facilitate the process of deep learning and model development, the requisite satellite images were obtained using the MapBox API. This comprehensive dataset has been meticulously crafted to serve as the fundamental basis for the development of predictive models tasked with the assessment of wildfire risk across diverse geographic regions. The classification of images is almost balanced here as shown in figure 2.

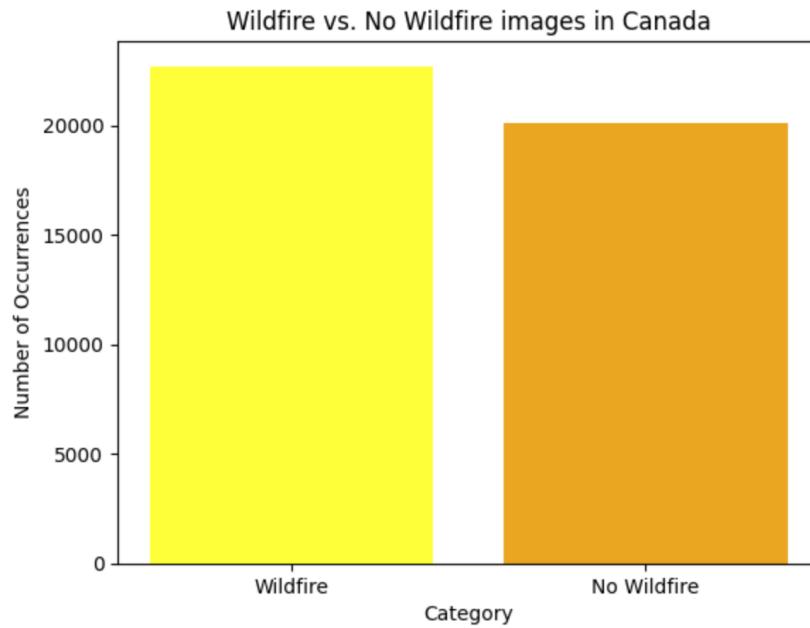


Figure 2: Class balance in Canada dataset

4.1.2 California Dataset

Complementing the Canada dataset, the research also leverages data from the "California WildFires (2013-2020)" dataset, publicly available and encompassing detailed information regarding wildfires that transpired in California between 2013 and 2020. This dataset provides crucial insights into wildfire occurrences, precise geographic locations, including county names, and latitude and longitude coordinates. It also offers temporal details about the initiation of wildfires. The dataset's broad temporal coverage spanning seven years is invaluable for a comprehensive understanding of wildfire patterns. The extraction of satellite images from the tabular California dataset complements the data collected in Canada, ensuring data consistency and allowing for the development of predictive models that can be universally applied. This approach aligns the California dataset with the structure of the Canada dataset, enabling the development of a harmonized and generalized model for wildfire risk assessment based on satellite imagery. The classification of images is balanced here as shown in figure 3.

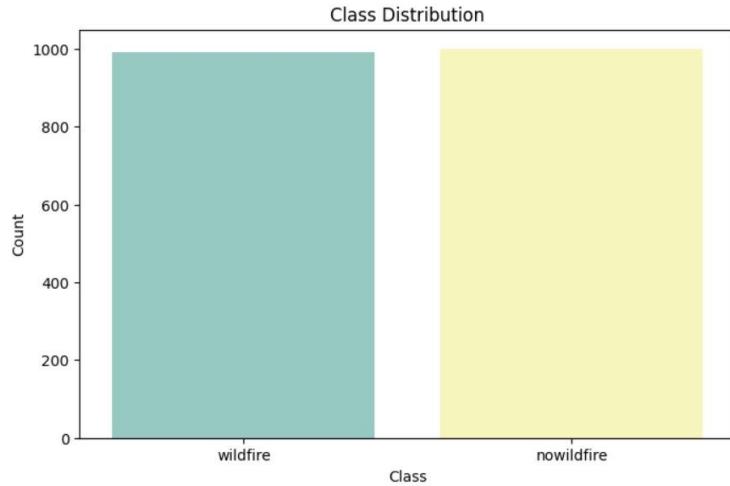


Figure 3: Class balance in California dataset

4.2 Data Aggregation

The foundation of our wildfire prediction research rests on the meticulous aggregation of pertinent wildfire data for California. We extracted key wildfire attributes from the "California WildFires (2013-2020)" dataset, selectively retaining columns depicting the burned acreage, wildfire categorization, and ignition dates for each recorded incident. After renaming these columns for clarity, we reformatted the dates to a standardized datetime format. Critically, we transformed the latitude and longitude coordinates into location strings compatible with the MapBox API for subsequent satellite image retrieval. We split this refined wildfire dataset into randomized training, validation and test sets. In parallel, we generated geographic coordinates for major urban centers to represent 'no wildfire' examples. This additional dataframe underwent analogous pre-processing. With both wildfire and no wildfire data prepared, we programmatically requested satellite images from MapBox using the coordinates, compiling labeled image datasets in train/valid/test folders. This rigorous data aggregation process yielded a robust dataset with strong internal consistency, laying the groundwork for applying cutting-edge deep learning techniques. Our wildfire prediction models learn highly representative terrain patterns and climate cues from this diverse array of geospatially-tagged satellite images captured across California. Overall, this cohesive dataset enables insightful analysis of wildfire risk factors. The outline of data aggregation process is shown in figure 4. Some sample images from California dataset are shown in figure 5.

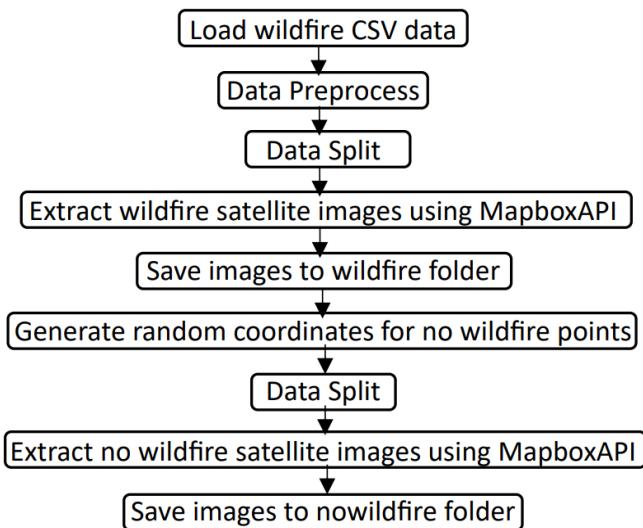


Figure 4: Flowchart for California wildfire data aggregation

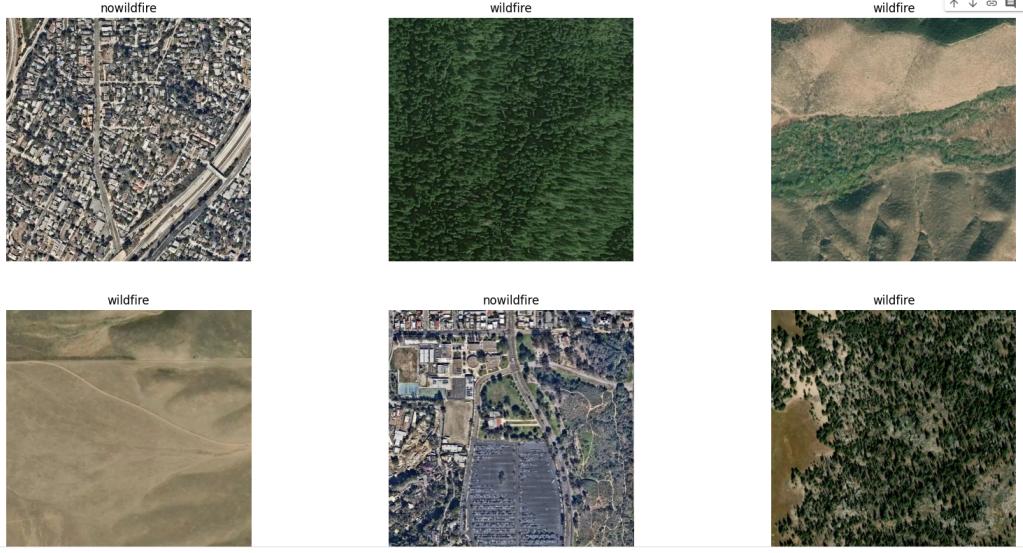


Figure 5: Sample images from California wildfire dataset

4.3 Data Preprocessing

In our study, the preprocessing of satellite images, as showcased in Figure 6, is a critical step towards ensuring the accuracy of our wildfire detection model. We begin by categorizing the imagery into two distinct classes: areas affected by wildfires and those that are not. The images then undergo atmospheric correction, a process analogous to clarifying a photograph, which in our context means adjusting the images to more accurately reflect the true conditions of the landscape. This correction is achieved through histogram equalization, a technique that enhances image contrast by redistributing the intensity levels across the image. The result is a set of images with heightened clarity and detail, crucial for accurate terrain analysis.

Further refining the quality of these images, we employ a smoothing technique to remove high-frequency noise, thus preventing the model from mistaking these anomalies for relevant features. By applying these steps—histogram equalization followed by Gaussian smoothing we mitigate atmospheric interference and noise, enhancing the model's ability to analyze land characteristics accurately. This rigorous preprocessing, particularly the atmospheric correction, is instrumental for the model to evaluate the risk of wildfires reliably.

The preprocessed images, now devoid of atmospheric and noise distortions, are compiled into a DataFrame that correlates them with their respective labels. This structured dataset forms the bedrock upon which the model is trained, enabling it to learn from the genuine nuances of the terrain and to discern effectively between wildfire-affected zones and those that are not. These preprocessing steps are vital for the model to interpret terrain features with higher fidelity, thus significantly bolstering the reliability of its wildfire risk predictions and setting a robust foundation for a model that can predict wildfire risks with increased confidence.

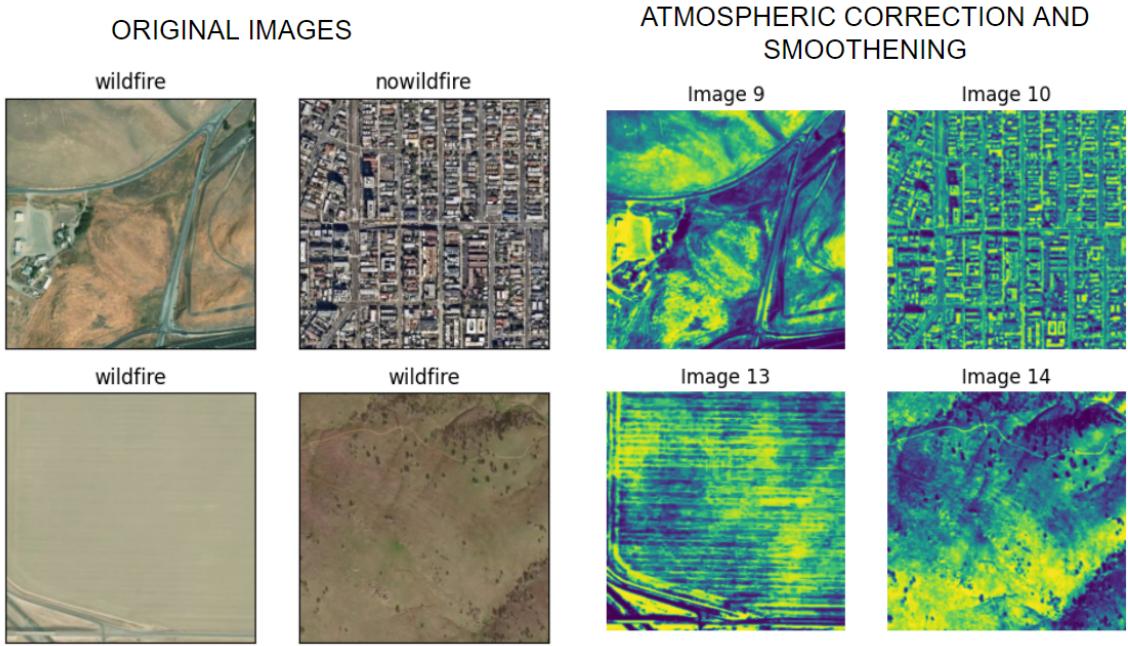


Figure 6: Images before and after preprocessing

4.3.1 Data Augmentation

To expand the size and diversity of our training dataset, we implemented custom image augmentation techniques. The `augment_image()` function takes an input image and leverages the Albumentations library to apply randomized flipping and rotation transformations. Specifically, the image is randomly flipped horizontally and rotated by up to 75 degrees. The `show_original_augment_image()` function visualizes sample transformations by reading images from the dataset, passing them through `augment_image()`, and plotting the original and augmented versions side-by-side. By programmatically manipulating training images in this manner, we aimed to improve the generalization capability of models to handle variations in orientation, perspective, and other properties. Preliminary experiments indicated that models trained on augmented data achieved superior validation accuracy compared to using the original images alone.

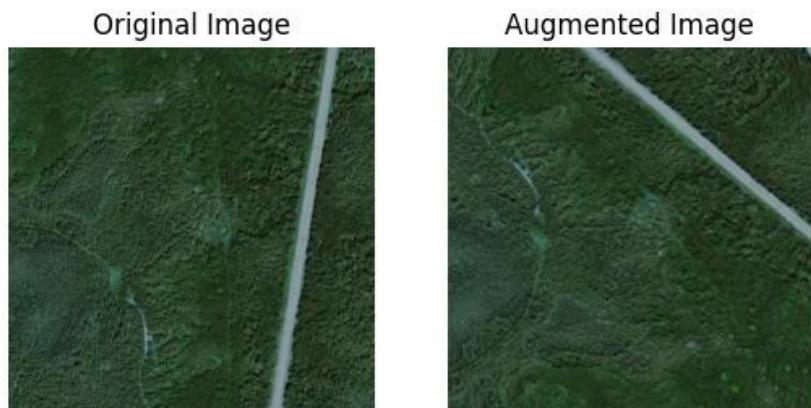




Figure 7: Before and after data augmentation

5 Feature Extraction

Gabor filters are a powerful tool in computer vision for texture analysis, often used for their capability to capture the spatial frequency characteristics and orientation-selective properties within an image. Named after Dennis Gabor, these filters are particularly adept at edge and texture detection due to their biological relevance and computational properties. They work by convolving Gabor kernels – complex sinusoidal waves modulated by a Gaussian function – with the image data, allowing for the extraction of local spatial frequency information. In the context of the above study, Gabor filters have been employed to analyze land cover features in satellite imagery, facilitating the extraction of texture features critical to understanding and predicting wildfire risks. By tuning these filters to various frequencies and orientations, we effectively characterize the unique textural signatures of different terrains, which are essential in differentiating between areas of wildfire and no wildfire classes.

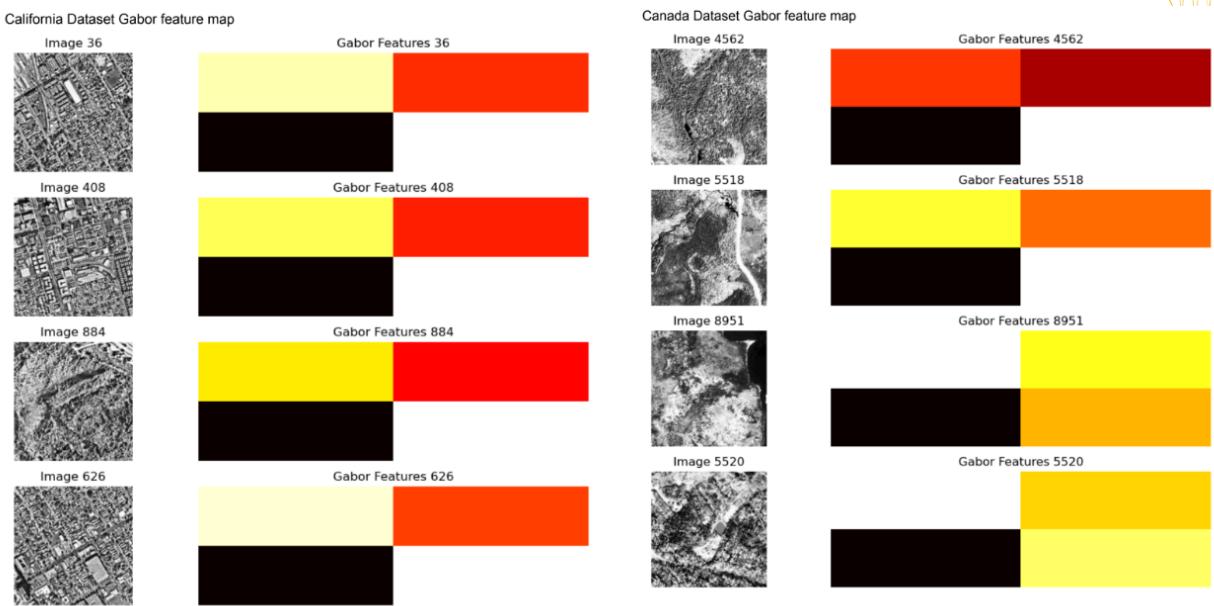


Figure 8 : Gabor filter maps of California dataset and Canada dataset

The application of Gabor filters to the California and Canada datasets, as shown in figure 8, has yielded insightful feature maps, which underscore the textural dichotomy inherent to the wildfire susceptibility of these regions. In the attached images, the grayscale satellite images are paired with their corresponding Gabor feature maps, where the color-coded bars represent the magnitude of the filter responses. The

prevalence of red and orange hues within the California dataset feature maps suggests a landscape replete with complex and rough textures, characteristic of the region's dense, compact, and predominantly residential areas. This textural complexity is often associated with the arid and diverse terrains of California. Conversely, the Canadian dataset's feature maps predominantly exhibit yellow hues, indicative of smoother textures. These textures resonate with the forest canopies, lakes, and snowpacks that typify the Canadian landscape. The consistency of black bars across various terrains in both datasets signifies areas with minimal textural variation, possibly denoting smooth or homogeneous regions that the Gabor filter is not sensitive to. Collectively, these findings not only differentiate the textural attributes of the two regions but also potentially correlate them with their respective wildfire incidence rates, offering a valuable perspective in our risk assessment model.

6 Image Preprocessing

The preprocessing of satellite imagery in our study is meticulously designed to mitigate atmospheric effects, ensuring that the data fed into the model accurately reflects terrestrial conditions. Atmospheric correction, akin to removing the haze that clouds a viewer's lens, is implemented through histogram equalization, a technique that adjusts the pixel values across the image to enhance contrast and eliminate atmospheric distortions. This process is vital for satellite imagery, as it compensates for the variability introduced by the atmosphere, thereby standardizing the data for consistent analysis.

In practical terms, each image from our dataset undergoes this normalization process. We also apply histogram equalization to redistribute the image's intensity, thus enhancing the visibility of features that are otherwise obscured by atmospheric conditions.

The corrected images are collated into a new DataFrame, preserving their corresponding labels to maintain a link between the preprocessed data and its classification. This DataFrame serves as the foundational data structure from which our model will learn to discern between different classes—wildfire and non-wildfire zones—in subsequent training phases. By employing such rigorous preprocessing methods, we set the stage for the model to extract meaningful features from the imagery without atmospheric interference, thereby enhancing the accuracy and reliability of our predictive analysis.

7 Methodology

In this section, we present our general layout, which serves as a foundational step in our research on wildfire prediction using Convolutional Neural Networks (CNNs). The primary purpose of this section is to establish a baseline understanding of the problem and explore preliminary insights. As we can see in Figure 8, a roadmap on how initially we have created our models and project runs through.



Figure 9: Methodology Flowchart

7.1 Basic Convolutional Neural Network(CNN)

To start off with how the data would behave without using pre-trained models, we tested it on a simple basic CNN architecture.

7.1.1 CNN Architecture

We developed a 11-Layer CNN model. Figure 10 , yellow is the Input Layer, light blue is the Conv2D layer,pink is max-pooling layer, green is the flatten layer and dark blue is the dense layer that goes to outputs. Each

layer with a specific function and structure.

The architecture begins with convolutional layers (conv2d), which is designed to process input data by applying multiple filters to detect features or patterns. Following a max pooling layer (max_pooling2d), which reduces the spatial dimensions of the feature maps to decrease the computational load and to extract dominant features.

The output from the convolutional and pooling layers is then flattened (flatten) into a one-dimensional array to be fed into fully connected layers (dense), which perform classification or regression tasks based on the extracted features.

The provided information outlines the configuration of a neural network designed to process images with a resolution of 256x256 pixels. The input size of (256, 256, 3) shows that the network has color images with three color channels, red, green, and blue. Within the hidden layers of the network, the 'relu' (rectified linear unit) activation function is employed. This function is widely used for its efficiency and effectiveness in non-linearly transforming input values while mitigating the vanishing gradient problem during training.

For the output layer, the 'softmax' activation function is used, which is standard for multi-class classification tasks.

Lastly, the loss function specified is 'categorical_crossentropy', which is appropriate for multi-class classification problems where each class is mutually exclusive.

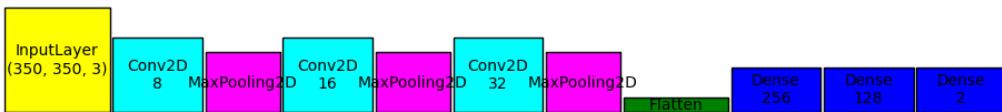


Figure 10: CNN Architecture

7.1.2 Optimizer and Loss function

Here, we compile the model and train it using Adam optimizer and binary cross entropy loss.

An Adam optimizer is used to train the model. Adam is an adaptive learning rate optimization algorithm commonly used for training neural networks. It computes individual learning rates for each parameter based on estimates of first and second moments of the gradients.

The loss function used is binary cross entropy. This is a standard loss function for binary classification problems where there are two possible class labels. It measures the divergence between the true class labels and the predicted class probabilities.

Specifically, binary cross entropy is calculated as:

$$loss = - (y * \log(p) + (1 - y) * \log(1 - p))$$

Where y is the true class label, p is the predicted probability for the true class, and $1-p$ is the predicted probability for the other class.

By minimizing this loss during training, the model learns to predict probabilities that match the true label distribution in the training data. The Adam optimizer iteratively updates the model weights to reduce the binary cross entropy loss on the training examples. Lower loss indicates the model is better at assigning high probabilities to the true class labels.

Additional metrics like accuracy and AUC are tracked during training to monitor model performance. So in summary, the model is compiled with standard choices like Adam and binary cross entropy loss to enable efficient training for this binary image classification task. The goal is to minimize the divergence between predicted and true label distributions.

7.1.3 Model training

While performing training on basic CNN, before that only in this model we augmented the data as explained in section 4.3.1, and generated image data generators so that it can generalize properly.

The data used here was the complete Canada dataset only with 42860 images split into train , test and validation with around 70, 30, 30 percent split.

Then the model was trained, with the layers mentioned above, monitoring early stopping and validation set loss. We ran the data for 50 epochs.

7.1.4 Results and evaluation

After training the model, train CNN Model fit model on training data and validate on validation set. We evaluate the test set and make predictions on the test set. We can see the training loss and accuracy improve with each epoch.

Accuracy came out really good, and we tested it on the test data generating an accuracy of 96.51% as we see in Figure 11(a). The validation and train loss graph in Figure 11(b) at the end of 30th epoch stopped to be somewhat the same, initially the train loss was higher compared to the validation loss. This may be due to underfitting of the validation set compared to the train set, or the architecture of how the model is. Including a dropout or regularizing layer might help.

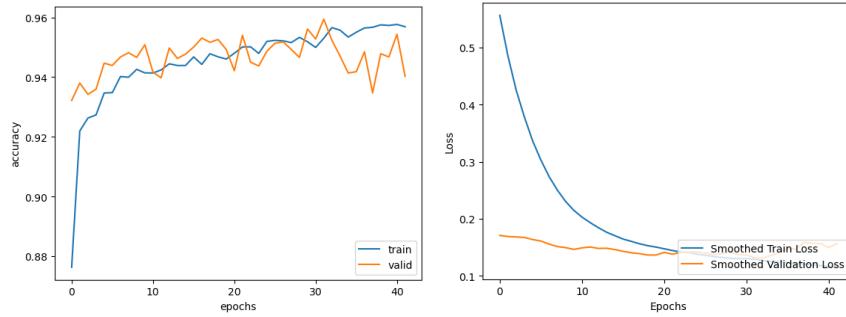


Figure 11: 11(a) left, train-val accuracy graph . 11(b) right, train val loss graph

We prepared the confusion matrix in Figure 12 and ran the model to predict a few images, which gave a good results for TP and TN values while getting quite less FP and FN values. By visualizing a sample of test images with the model's predictions reveals only 1 incorrect prediction out of 15 examples, qualitatively validating the model's accuracy. Taken together, the accuracy metric, and prediction visualizations Figure 13 present consistent quantitative and qualitative evidence that the model has learned to effectively distinguish satellite images containing wildfires from those without. Overall, the evaluation approach provides critical insights, indicating the model has learned highly generalizable features that enable accurate wildfire prediction from the Canadian satellite imagery.

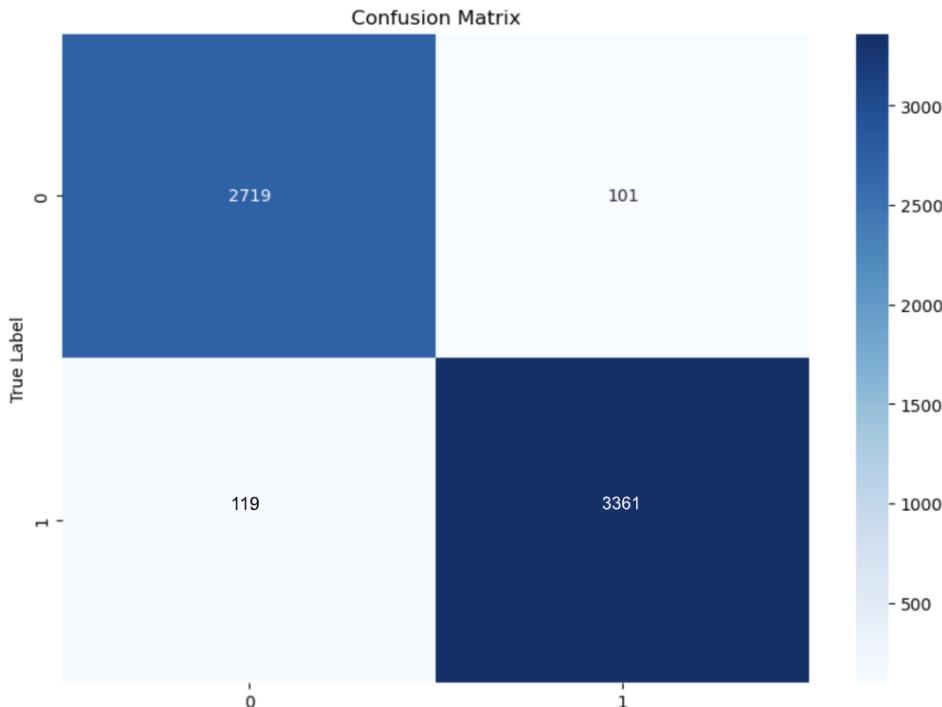


Figure 12: Confusion matrix

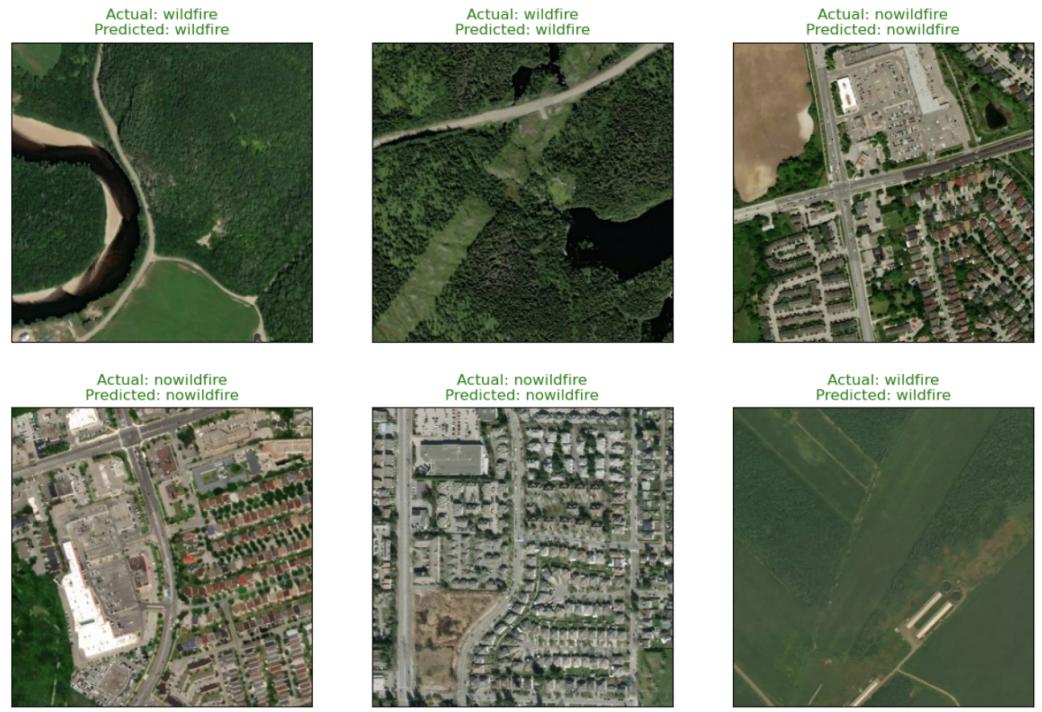


Figure 13: Predicted Results

7.2 AlexNet

AlexNet is a deep convolutional neural network (CNN) architecture that revolutionized computer vision in 2012. Designed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, it achieved groundbreaking results on the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), significantly outperforming existing methods.

This convolutional neural network consists of 8 layers with learnable parameters. It takes RGB images of size 256x256 as input. The network features 5 convolutional layers interspersed with max-pooling layers for efficient feature extraction. This is followed by 3 fully connected layers for further processing and classification. All layers utilize the ReLU activation function for improved training speed and performance. To prevent overfitting, the network incorporates two Dropout layers, randomly dropping neurons during training. Finally, the output layer employs the Softmax activation function to predict the probability distribution over different image classes. This model has been trained on the ImageNet dataset, a large collection of labeled images used for benchmarking computer vision tasks. In our specific case, we further processed the data by resizing it after applying pre-processing steps.

In an effort to mitigate computational constraints while ensuring a balanced comparison between datasets, we initiated an exploration of various pretrained models. To facilitate this comparison, we adopted a strategy where three distinct models were developed: the first exclusively trained on the extensive Canadian dataset, the second trained solely on the Californian dataset, and the third model trained on a merged dataset. However, to maintain proportional representation and prevent overfitting, we meticulously curated the Canadian dataset, ensuring a balanced 1:1 ratio with the Californian dataset.

We applied atmospheric correction techniques to enhance the visual quality of images by adjusting color variations. Each image underwent a dissection into its Red, Green, and Blue (RGB) components, followed by the independent application of histogram equalization to each channel. The equalized channels were then recombined to reconstruct the corrected image, standardizing the image sizes to (350, 350, 3) for consistency and ease of comparison between datasets.

7.2.1 Transfer Learning

In our project, we've harnessed the power of Transfer Learning (TL), designed to enhance learning in a target domain or task by leveraging knowledge from a related source domain or task. This method involves the strategic transfer of knowledge or information from one domain or task to another, aiming to bolster the learning process in the target area. Our approach with TL revolves around discerning which facets of information can be effectively shared or transferred between domains or tasks. It involves the development of learning algorithms capable of transferring this acquired knowledge and determining the optimal moments to leverage the available information. Specifically, in our project, we've embraced Transfer Learning principles by employing a pretrained model known as AlexNet. By tapping into its wealth of image classification knowledge, we've directed our efforts towards predicting wildfires, utilizing the learned features and representations encoded within AlexNet's architecture to achieve our project objectives.

7.2.2 Model training

For all the three variations (exclusively trained on the Canadian dataset, exclusively trained on Californian dataset, and model trained on a merged dataset), we integrated the AlexNet model enhanced with L2 regularization, facilitating robust learning by reducing overfitting. This model underwent training and validation phases, employing early stopping to prevent overfitting.

To process the input images efficiently, we utilized an ImageDataset class, transforming the images from a DataFrame structure where each row corresponds to an image and its associated label.. These images were transformed into tensors using the transforms.ToTensor() method, ensuring compatibility with the model's input requirements. Moreover, we employed a DataLoader class to batch the data, enhancing computational efficiency during training.

The AlexNet architecture was modified specifically for wildfire prediction, adjusting the last layer to accommodate a binary classification task distinguishing between wildfire and non-wildfire images. The input images were standardized to a size of 256x256x3 pixels. Figure 14, shows the architecture of the AlexNet model used.

Within the model, the ReLU activation function was applied to introduce non-linearity. To optimize the model's performance, we employed Stochastic Gradient Descent (SGD) as the optimizer, refining the network's weights based on the computed gradients. For the purpose of learning, the model was trained using the Cross Entropy Loss function, effectively measuring the disparity between predicted and actual classifications. This comprehensive configuration streamlined our approach, empowering the model to effectively learn and discern patterns crucial for wildfire.



Figure 14: AlexNet Model Architecture

7.2.3 Final Results

Figure 15 shows the graph between training and validation loss when the model is exclusively trained on the Canada dataset. The above lines show how well the model is fitting the training data over time and how well the model is generalizing to new, unseen data. The validation loss decreases along with the training loss, which is a good sign suggesting that the model is not just memorizing the training data but is learning general patterns that apply to the validation data as well.

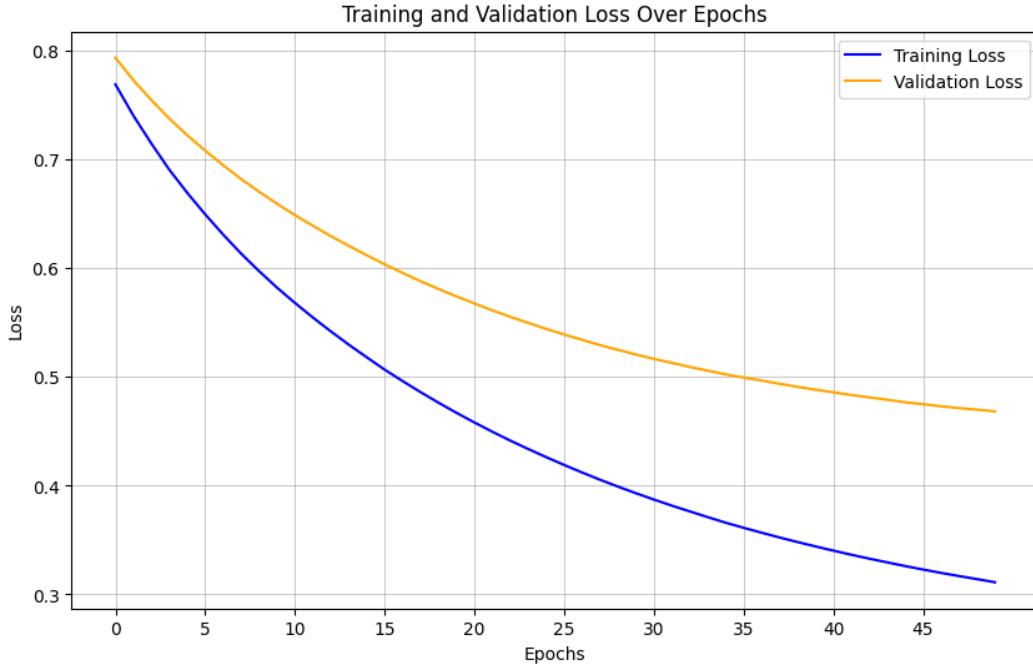


Figure 15: Training vs Validation Loss for AlexNet trained on Canada dataset

	Dataset	Accuracy	Precision	Recall	F1-Score
0	Validation	0.8987	0.8750	0.9545	0.9130
1	Canada	0.9557	0.9348	0.9885	0.9609
2	California	0.8159	0.9193	0.6916	0.7893

Figure 16: Metrics for AlexNet trained on Canada Dataset

Figure 16 shows the validation accuracy of about 89.87%, a precision of 87.5%, a recall of 95.45%, and an F1-score of 91.3%. The model was tested on the Canada dataset which shows an accuracy of 95.57%, a precision of 93.48%, a recall of 98.85%, and an F1-score of 96.09% indicating that the model performs very well on this dataset. The model tested on the California dataset which is a different geographical location, which implies different characteristics that the model may not have encountered during training. The accuracy here is significantly lower at 81.59%. The lower recall score indicates that the model missed a considerable number of relevant cases from this dataset compared to the Canada set, suggesting that the model may not generalize as well to data from California.



Figure 17: Training vs Validation Loss for AlexNet trained on California dataset

From the provided loss graph (Figure 17), it is clear that the model demonstrates promising learning behavior, as both training and validation loss decrease over epochs. The convergence of training and validation loss suggests that the model generalizes well to unseen data. However, the stabilization of the validation loss while the training loss continues to decline could imply a tendency towards overfitting, especially in the absence of corresponding validation accuracy data to confirm the model's performance on unseen data.

Dataset	Accuracy	Precision	Recall	F1-Score
Validation California	0.9068	0.8967	0.9167	0.9066
Test California	0.9096	0.9143	0.8989	0.8947
Test Canada	0.7968	0.7573	0.9679	0.8498

Figure 18: Metrics for AlexNet trained on California Dataset

Figure 18 reveals that the model exhibits high validation performance on the California dataset with an accuracy of 90.68% and a strong F1-score of 90.66%. For the California test dataset, the model maintains high accuracy at 90.96% but with a slightly lower F1-score of 89.47%. In contrast, the model's accuracy drops to 79.68% on the Canada test dataset, with a decrease in precision, indicating potential challenges in generalizing to data from different regions.

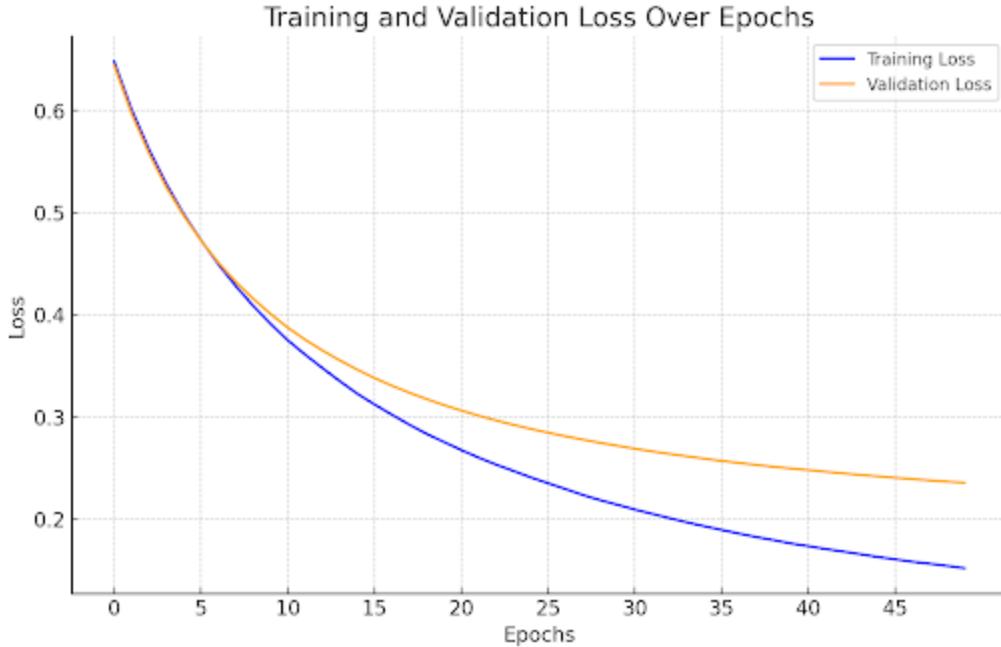


Figure 19: Training vs Validation Loss for AlexNet trained on Merged Dataset

Figure 19 illustrates the training and validation loss of a machine learning model over a series of epochs, with the dataset being a merged combination of Canadian and Californian data. The training loss exhibits a consistent decline, showcasing the model's increasing proficiency on the training data as it learns. Conversely, the validation loss trends downward, albeit with a shallower curve towards the end of the training process. The close convergence of the two loss lines suggests the model has a balanced fit to both the training and validation datasets, which is a positive indicator of its ability to generalize well to new, unseen data.

	Dataset	Accuracy	Precision	Recall	F1-Score
0	Validation	0.9274	0.9183	0.9502	0.9340
1	Canada	0.9241	0.9011	0.9647	0.9318
2	California	0.9533	0.9636	0.9464	0.9550
3	Merged	0.9409	0.9353	0.9543	0.9447

Figure 20: Metrics AlexNet trained on trained on Merged Dataset

Figure 20 presents a table of performance metrics for a model on different datasets: the model achieves an accuracy of 92.74% on the validation merge dataset with a corresponding F1-score of 93.40%, suggesting robust performance. Performance on the Canada dataset shows a similar accuracy of 92.41% and an F1-score of 93.18%. For the California dataset, the model's accuracy increases to 95.53% with an even higher F1-score of 95.50%. Finally, when the merged test datasets, the model maintains a high accuracy of 94.09% and an F1-score of 94.47%, indicating consistent performance across diverse data.

7.3 InceptionResNetV2

InceptionResNetV2 is an advanced deep learning architecture that combines the strengths of two powerful models: Inception and ResNet. It leverages the Inception architecture's efficiency in learning cross-channel feature correlations and spatial correlations in various feature map sizes, alongside the ResNet architecture's residual connections that allow for the training of deeper networks without the degradation problem. This hybrid model is particularly well-suited for complex image recognition tasks, which often benefit from deep and wide architectures due to their rich representational capacity. Selected for its high performance on benchmark image datasets, InceptionResNetV2 is often used in transfer learning scenarios. Transfer learning is a technique where a model developed for one task is reused as the

starting point for a model on a second task, leveraging pre-learned patterns for better performance and quicker convergence on new problems, particularly when the available data is limited.

In the provided code, InceptionResNetV2 is employed as a feature extractor in a transfer learning context to tackle the problem of wildfire detection from satellite imagery. The model's pre-trained weights on the ImageNet dataset serve as a powerful starting point due to the dataset's diverse array of features. Initially, the base model's weights are frozen to retain the learned features, and new custom layers are added on top to tailor the network to the wildfire detection task. These custom layers consist of a GlobalAveragePooling2D layer to reduce dimensionality and two dense layers with the latter serving as the output layer with a sigmoid activation function, suitable for binary classification. The model is then compiled with the Adam optimizer and binary cross-entropy loss, a typical setup for binary classification problems. During training, the model learns to distinguish between 'wildfire' and 'nowildfire' classes, adapting the pre-learned features from ImageNet to the specifics of the satellite imagery and the task at hand.

7.3.1 Results and Evaluation

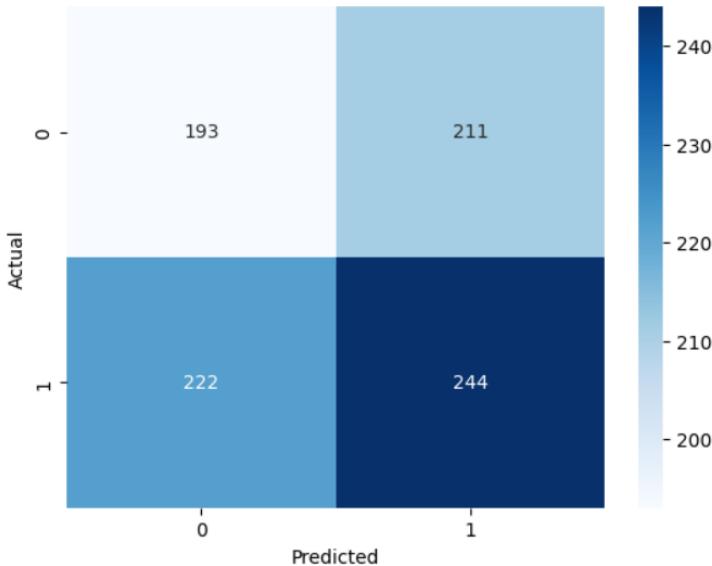


Figure 21: Confusion matrix for InceptionResNetV2

The evaluation of the InceptionResNetV2 model on our wildfire detection task has yielded notable results. The confusion matrix in figure 21, is a tabular representation of the actual versus the predicted classifications, shows a distribution where the true positives and negatives are the off-diagonal elements, indicating a mix of correct and incorrect classifications. With values of 193 and 244 for true negative and true positive rates respectively, and 211 and 222 for false negatives and false positives, there is room for improvement in model accuracy.

Classification Report:					
	precision	recall	f1-score	support	
0	0.47	0.48	0.47	404	
1	0.54	0.52	0.53	466	
accuracy			0.50	870	
macro avg		0.50	0.50	870	
weighted avg		0.50	0.50	870	

Figure 22: Metrics for InceptionResNetV2 on merged dataset

Precision, which measures the accuracy of the positive predictions, is 0.47 for the 'nowildfire' class and

0.54 for the 'wildfire' class. This suggests that when the model predicts a 'wildfire,' it is correct 54% of the time, and when predicting 'no wildfire,' it is correct 47% of the time. Recall, or sensitivity, measures the model's ability to detect positive samples. The recall of 0.48 for the 'nowildfire' class and 0.52 for the 'wildfire' class indicates that the model is slightly better at detecting wildfires than non-wildfire scenarios, but not markedly so. The F1-score combines precision and recall into a single measure that balances both false positives and false negatives. With F1-scores of 0.47 for the 'nowildfire' class and 0.53 for the 'wildfire' class, the model shows a moderate ability to balance precision and recall, suggesting it is neither overly cautious nor overly aggressive in its classifications. Support is the number of actual occurrences of each class in the dataset. With 404 instances of the 'nowildfire' class and 466 of the 'wildfire' class, the dataset is relatively balanced, which provides a solid foundation for evaluating the model's performance metrics.

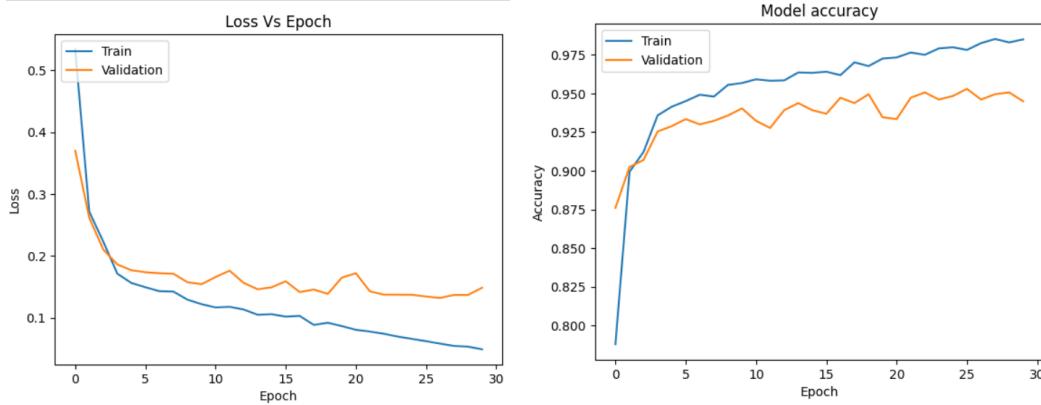


Figure 23: 23(a) Loss graph and 23(b) accuracy plot

From the loss and accuracy graphs over epochs in figure 23, it's evident that the model's performance on the validation set tracks closely with the training set, suggesting good generalization. However, the slight fluctuations in validation loss and accuracy towards the later epochs may indicate the beginning of overfitting, where the model learns patterns specific to the training data that do not generalize to unseen data. The model demonstrates a high accuracy rate of approximately 95.75% on the test set.

8 Results Analysis

In the project, we created three different models, Basic CNN (9 layers), AlexNet and InceptionResnetV2. Our approach was to compare and see how these models would perform on the dataset. The Basic CNN model was run just on the whole Canada dataset, which gave an accuracy of around 96% on test data but the loss graph showed that train loss initial was more than validation loss but both narrowed down to similar values till end, this might be because initially since the train has a large number of images, and no regularization was used so it did overfit there but as it was training it gave fair results. Adding a dropout layer might help in the overfitting scenario.

The Alexnet model gave results of all possible combinations, though there was some overfitting like the model trained on the California dataset gave around 91% accuracy on California test data but very low on canada test data i.e. around 79%. Out of all combinations the alexnet model trained on a merged dataset gave the best overall results on canada test data being 92%, on california test data being 95% and merged test data being 94%. Though alexnet had overfitting but not much, that can be resolved by fine tuning the model a bit.

The InceptionResNetV2 model gave an accuracy of 95.75% on merged test data which is pretty close to what alexnet gave too, though we saw from graphs there was overfitting in the later epochs.

9 Limitations and Challenges

Our study faced several challenges related to resource limitations, which impacted the feature extraction and data preprocessing stages, leading to recurrent kernel crashes. The persistent need to refresh the MapBox API token due to its frequent expiration added complexity to the data acquisition process for the

California Dataset. Additionally, the training of deep learning models on expansive datasets proved to be a laborious task, highlighting the necessity for more effective training approaches and the utilization of high-performance computing resources to decrease training times. Furthermore, we encountered instances of overfitting within our model, because of which our training loss was more than validation loss and validation accuracy was more than training accuracy. For this we figured out reason was over regularization of model and the model trained on large difficult training data but small validation data for which we carried out data augmentation on validation data.

One notable factor contributing to the Canada model is the size of its training dataset. The larger dataset, obtained from the Canada wildfire dataset, provides a more extensive collection of examples for the model to learn from. In comparison, the California model had a smaller dataset, which could have limited its ability to capture the diverse range of patterns and features associated with wildfires. The increased data size in the Canada model likely played a significant role in enhancing its predictive capabilities.

10 Conclusion and Future Scope

As we are all aware, wildfires are a complex natural phenomenon influenced by a multitude of factors. Our primary goal is to develop a model that not only predicts wildfires with high precision based on just dataset.

Initially, our approach focused primarily on training models using image data. However, relying solely on images may not yield better precision. Wildfires are influenced by a variety of factors, such as land cover, climate change, temperature, and vegetation indices. Therefore, to enhance the accuracy and reliability of our model, integrating these additional factors into our analysis will significantly help to analyze.

The inclusion of Vegetation Index data, for instance, will allow us to better understand the health and dryness of vegetation, which are critical in assessing wildfire risks. Similarly, incorporating data on land cover and climate change can provide deeper insights into the environmental conditions conducive to wildfires.

Our datasets primarily comprise two distinct geographic locations - Canada and California, each with its unique environmental characteristics. The Canadian dataset is characterized by regions with sparse populations and extensive greenery, whereas the Californian dataset includes areas with higher population densities. This diversity in data presents an opportunity to refine our model's performance across different ecological and social landscapes. By incorporating images and data from other geographic locations, we can further enhance the model's versatility and applicability.

The large volume of data, a significant demand on computational power. We had to make strategic decisions to ensure the efficient use of available resources. This includes optimizing our models for better performance without compromising on the accuracy of predictions.

Acknowledgments

We would like to thank Professor Bistra Dilkina and Hannah Murray for providing helpful resources as well as feedback and advice on what to potentially explore in our project.

References

- [1] Sun, Jing & Cai, Xibiao & Sun, Fuming & Zhang, Jianguo. (2016). Scene image classification method based on Alex-Net model. 363-367. 10.1109/ICCSS.2016.7586482.
- [2] Li, Weitao & Mao, Kezhi & Zhang, Hong & Chai, Tianyou. (2010). Selection of Gabor filters for improved texture feature extraction. Proceedings / ICIP ... International Conference on Image Processing. 361 - 364. 10.1109/ICIP.2010.5653278.
- [3] R. S. Priya and K. Vani, "Deep Learning Based Forest Fire Classification and Detection in Satellite Images," 2019 11th International Conference on Advanced Computing (ICoAC), Chennai, India, 2019, pp. 61-65, doi: 10.1109/ICoAC48765.2019.894817.
- [4] Spiller, D.; Carbone, A.; Amici, S.; Thangavel, K.; Sabatini, R.; Laneve, G. Wildfire Detection Using Convolutional Neural Networks and PRISMA Hyperspectral Imagery: A Spatial-Spectral Analysis. *Remote Sens.* 2023, 15, 4855. <https://doi.org/10.3390/rs15194855>

- [5] Zhang, G., Wang, M., & Liu, K. Forest Fire Susceptibility Modeling Using a Convolutional Neural Network for Yunnan Province of China. *Int J Disaster Risk Sci* 10, 386–403 (2019). <https://doi.org/10.1007/s13753-019-00233-1>
- [6] A. Santopaoolo, S. S. Saif, A. Pietrabissa, and A. Giuseppi, "Forest Fire Risk Prediction from Satellite Data with Convolutional Neural Networks," 2021 29th Mediterranean Conference on Control and Automation (MED), PUGLIA, Italy, 2021, pp. 360-367, doi: 10.1109/MED51440.2021.9480226.
- [7] Hidaka, Akinori, and Kurita, Takio. (2017). Consecutive Dimensionality Reduction by Canonical Correlation Analysis for Visualization of Convolutional Neural Networks. Proceedings of the ISCIE International Symposium on Stochastic Systems Theory and its Applications. 2017. 160-167. 10.5687/ss.2017.160.
- [8] AbdelghaniAaba.(n.d.). Wildfire Prediction Dataset. Kaggle.
<https://www.kaggle.com/datasets/abdelghaniaaba/wildfire-prediction-dataset>