Wildfire Detection from RBG Images using CNN, Transformer and Hybrid models
Intermediate Progress Report

Matiss M Mednis, Nathan Loh, Nikolaus Schultze
CS577 - Deep Learning
Illinois Institute of Technology

11/1/2023

Introduction and Problem Description

Wildfires are detrimental to ecosystems and humanity as a whole. They can be started for various reasons and spread very quickly. Recently with the Canadian wildfires, we have seen the global impact a large-scale forest fire can cause. Wildfires can spread rapidly and become extremely difficult to control. Forest agencies attempt to do their best to prevent them, but even the slightest initial spark can have massive consequences. Prevention is one half of the equation, and the other is detection. Early detection systems of wildfires can help prevent large-scale disasters and allow authorities to identify and respond quickly to wildfires before they become out of control. Various automated methods of monitoring wildfires exist such as inferred cameras, lasers, and classifying RGB video feeds or images. RGB images and video from monitoring stations above the treeline may be of the most interest to monitoring wildfires as RGB cameras are the most accessible and the cheapest. Satellite image detection may only be effective when fires are already at a large scale. Therefore, RGB images on the ground from monitoring stations throughout the most prone portions of a forest may be of interest for their potential early detection capabilities as well. Augmented with direct human monitoring, an automated system could monitor difficult to observe places for humans and thus expand the monitoring potential beyond what humans alone could practically achieve. Once potential fires are flagged, they can be further investigated by forest services and prevented earlier. Therefore, developing models that can most accurately identify the beginning and intermediate stages of a wildfire based on live RGB image feeds would be of great interest and importance in keeping ecosystems and forests intact and saving lives and property.

Challenges with wildfire detection based on purely colored images, is that naturally occurring phenomenon such as mist, clouds, and sunsets and sunrises can all prove difficult for image classification systems as under proper conditions, the colors and shapes produced by these occurrences can look similar to wildfires. In those cases, approaches such as satellite monitoring or infrared detection would likely perform better in reducing false positives. Satellite monitoring stations offer larger fields of view and an angle from above which may help cut through minor weather patterns such as mist or a very orange and red evening sunset or morning sunrise. Thermal monitoring offers information on the temperature, and thus also would be less likely to produce false positives where the sky or mist may give off an orange or red hue similar to that of a wildfire.

In our work, we apply CNN, Transformer, and Hybrid CNN Transformer models to a dataset of wildfire images and forest images to address the challenge of identifying wildfires with RGB images. We aim to compare the performance of different models within those three model architectures to produce the most effective model for wildfire observation and monitoring. We utlilize accuracy, recall, precision, F1, and loss in order to evaluate the effectiveness of the models. Additionally, we may add images of fires outside of wildfire contexts to increase the variation and diversity of the fire class to analyze the impact on performance or implement the Multi-Task Learning framework of the dataset - as outlined in the dataset description below - if time permits. These two additions may offer increased effectiveness of our proposed model and monitoring system.

Description of the Data

The wildfire dataset is a comprehensive collection that explores the effectiveness of RGB imagery for forest fire detection using machine learning techniques. It comprises 2,700 aerial and ground-based images from various online sources, including government databases, Flickr, and Unsplash. These images cover various environmental conditions, forest types, geographical locations, and the intricate dynamics of forest ecosystems and fire events, making them a valuable resource for forest fire detection research. All images are sourced from the Public Domain. Detailed information about the origin URLs and resolutions of each image is provided in the data source link. A unique feature of this dataset is utilizing a Multi-Task Learning framework designed to enhance forest fire detection by addressing multi-class confounding elements. This approach aims to improve model accuracy and reduce false alarms, especially compared to traditional classification techniques. We aim to potentially implement this framework, as time permits. The dataset is divided into three main directories: Training (70%), Validation (15%), and Testing (15%). The nofire folder contains 1653 images, and the fire folder contains 1047 images. Thus about 61% of the data is labeled nofire and 39% of the images are labeled fire. One of the challenges of wildfire data is there aren't as many available images of the fire class as obtaining images of wildfires is challenging and doesn't often occur. Within the nofire folder, there are 847 images containing Forested areas without confounding elements, 336 images of Fire confounding elements, and 471 images of Smoke confounding elements. Within the fire folder, there are 662 images of Smoke from fires and 384 images of both smoke and fire. This further segmentation of the data and confounding elements may be implemented in attempt to improve model performance. The original dataset can be found at the following source https://www.kaggle.com/datasets/elmadafri/the-wildfire-dataset.

What have we done

Currently, we have initial results for our CNN models. We developed both simple and complex CNNs, where our complex CNN has increased depth and complexity. For the pretrained model, we applied VGG-16. The results were obtained from 2 epochs, as we aim to reduce the image sizes and increase our computational power to allow for more reasonable training times so that we may run more epochs. In analyzing the loss and accuracy graphs, we see more training can be done as the validation accuracy and loss have not plateaued. Additionally, we have generated confusion matrices for our completed models on the test set.

Model	**Accuracy**	**F1**	**Precision**	**Recall**
Simple CNN	75.85%	0.82	0.75	0.92
Complex CNN	66.10%	0.77	0.66	0.92
VGG-16	84.39%	0.88	0.83	0.94

Table 1. Test set results of models on wildfire dataset after 3 epochs of training

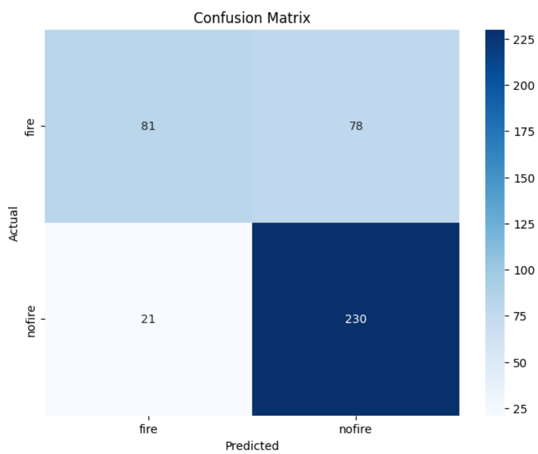


Figure 1. Confusion Matrix for Simple CNN

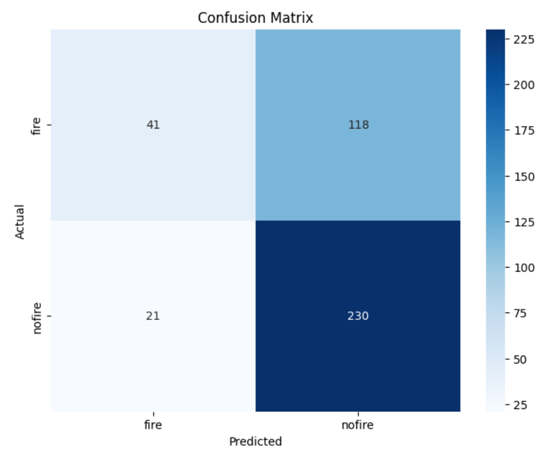


Figure 2. Confusion matrix for Complex CNN

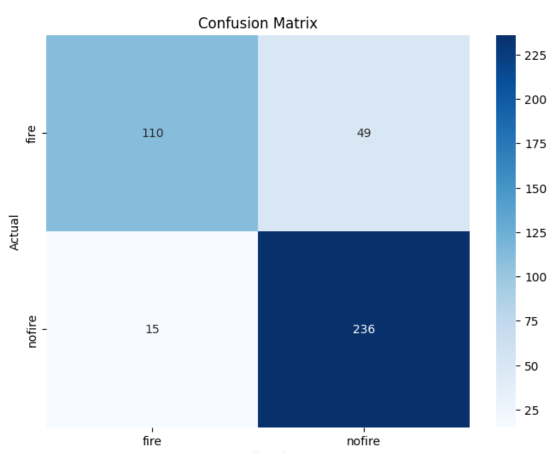
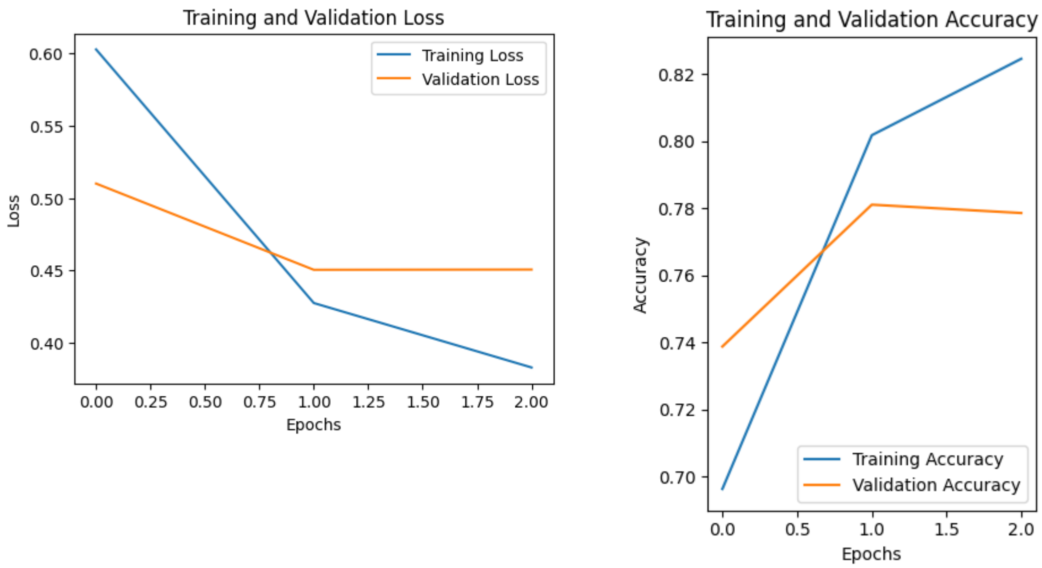


Figure 3. Confusion matrix for VGG-16



Figures 4 & 5: VVG-16 training and validation loss and accuracy

Remaining work

- Tune CNN Models: Refine the hyperparameters and architecture of the Convolutional Neural Network (CNN) models used for image classification. Fine-tuning aspects such as the number of layers, filter sizes, learning rates, and more to improve model performance.
- Pre-process Data to Reduce Training Time and Computational Requirements: Optimize the data preprocessing pipeline to reduce the time and computational resources required for training. This could include resizing images, normalizing data, or other techniques.
- Finish Developing Transformer Model by Hand (if time permits): Develop a Transformer model from scratch.
- Finish Implementing ViT Pretrained Model and Tune: Implement the Vision Transformer (ViT) model, a state-of-the-art architecture for image classification.
- Finish Developing Hybrid CNN+Transformer Model by Hand (if time permits): Explore a hybrid model that combines both CNN and Transformer architecture.
- Finish Implementing Pretrained Hybrid Model and Tune: If a pre-trained hybrid model exists, implement it and fine-tune it for your specific image classification task.
- Obtain Metrics for Models: Evaluate the models' performance by obtaining metrics such as accuracy, precision, recall, F1 score, and possibly more.
- Analysis: Analyze the results and metrics obtained from the models. Identify areas where models perform well and areas where they need improvement. This analysis can inform further model refinements.

- Append More Diverse Fire Images to Training Dataset (if time permits): Enhance the training dataset by adding more diverse fire images. A larger and more diverse dataset can help improve the model's ability to generalize to different scenarios.
- Final Report: Compile all the findings, results, and insights into a final report. This report can be used to communicate the project's outcomes, the performance of the models, and any potential recommendations for future work or improvements.

Github

https://github.com/nschultze/CS577Project