Wildfire detection of RGB images using CNN, Transformer models, and hybrid CNN and transformer models and the impact of performance by including fire images outside of wildfire context during training

Team Members: Matiss Mednis, Nathan Loh, Nikolaus Schultze

Description of the Problem

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 mistake of a dropped cigarette or poorly planned campfire 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 either satellites or monitoring stations above the treeline may be of the most interest to monitoring wildfires as RGB cameras are the most accessible and the cheapest. Therefore, developing models that can most accurately identify the beginning stages of a wildfire would be of great interest and importance in keeping ecosystems and forests intact and saving lives and property.

History of Previous Works

Most of the work on wildfire recognition systems is very new and published in 2023. Current work typically focuses on utilizing CNN MobileNetV2 models to detect wildfires [1] [2]. The data used comes from satellite images [2] [3] and in another case uses an on-the-ground, RGB set of images [1] with MobileNetV2. In other work, we see the use of transformer models on satellite images which produces better results than the CNN models [4]. In [1] they create a multi-level classification which first determines if there is smoke or not in an image and then further classifies whether there is a fire. This assisted in classifying images more accurately that have very red sunsets or mists in them. The separation of more granular classes was a good method that proved to do better than the binary class model. We propose training transformer models and hybrid CNN and transformer models on the on-the-ground RGB images of forest fires in various stages to improve the efficacy of camera monitoring stations in forests. We also propose utilizing images of fires outside of the forest fire setting, to improve the model's ability to recognize the traits of fire and smoke and potentially build a multi-task learning approach pipeline with additional training images of fires and smoke in general settings. By removing the

context of the forests, we hope the models might pick up on the more general traits of a fire and thus improve their efficacy in identifying the traits of fires in a forest. Essentially we consider if a more general fire detection system may transfer to forest fires and improve model performance as well as propose transformer model architecture in forest fire detection.

Milestones and Premilinary Plan

We first plan to replicate the CNN results as found in the paper [1] using the pre-trained MobileNetV2 CNN model. Then we plan to implement a transformer model and a hybrid CNN and transformer model - selecting from pre-trained models - and see if we have improvements in our wildfire detection to the CNN model. Finally, we explore augmenting the dataset by including images of fires with and without smoke and smoke without and without fires in an attempt to increase the variety of fire images the classifier is exposed to and measure the impact of this additional, related data on the model's performance. This idea is similar to that found in the paper [1]. So, we will combine datasets with images of fire labeled 'fire' along with the original wildfire dataset to see if the models can better capture the qualities of a fire - which are consistent among all instances of fires i.e. smoke and flames - by separating the images from a forest context.

October 11, 2023 - Apply 2 CNN Algorithms and Tune -
October 18, 2023 - Apply 2 Transformer Algorithms and Tune
October 25, 2023 - Apply 2 Transformer/CNN hybrid models and tune
November 1, 2023 - Intermediate Project Report Due
November 10, 2023 - Append general fire images and re-train/test models
November 8 - December 1, 2023 - Tentative Final Project Presentation
December 1, 2023 - Final Project Report Due

Github

https://github.com/nschultze/CS577Project

References

[1] https://www.mdpi.com/1999-4907/14/9/1697
[2] https://www.mdpi.com/2571-6255/6/4/169
[3] https://www.tandfonline.com/doi/full/10.1080/19475705.2023.2196370
[4] https://www.tandfonline.com/doi/full/10.1080/19475705.2023.2196370
https://www.kaggle.com/datasets/elmadafri/the-wildfire-dataset/data
https://www.kaggle.com/datasets/phylake1337/fire-dataset