Group 73 Final Report:
Wildfire Detection Classification Plan

Andy Huynh, Berk Yilmaz, Tanisha Tasnin
huynha3@mcmaster.ca, yilmag1@mcmaster.ca, tasnint@mcmaster.ca

1 Introduction

«««< HEAD Wildfires are becoming increasingly prevalent and destructive due climate change, particularly in northern and temperate forest regions such as Canada [1]. This poses a growing threat to both human safety and the environment with upto 2740 deaths in Canada between 2013 to 2018 alone [7]. The economic consequences are equally substantial: a good example being the 2020 Australian mega-fires which caused approximately US $20 billion in damages [15]. Wildfires also have profound and well-documented impacts on air quality and biodiversity worldwide [13]. Early detection of wildfires is crucial for effective response and mitigation efforts. However, detecting wildfires at an early stage can be challenging. Subtle indicators such as light smoke or small flames can be obscured by dense vegetation, clouds, or varying terrain, often making traditional monitoring methods unreliable [22]. In an attempt to address this issue, our project explores the use of Convolutional Neural Networks (CNNs) for automatic wildfire detection from images. Prior work by Tsalera et al. and Spiller et al. provided valuable insights for the research going into our project, highlighting both foundational and advanced CNN-based wildfire detection techniques [21], [16]. Our project formulates the task as a binary classification problem: given an RGB image of an area, the model identifies it as a “fire” or “no fire” scenario using ‘The wildfire dataset’ by El Madafri on kaggle [9]. To benchmark performance, we begin with a conventional CNN serving as our baseline. We then created 3 separate and individual models that used modern architectures: ResNet18, MobileNetV2, and EfficientNet-B0. These model choices were motivated by influential works from He et al., Tan et al., and Howard et al. [11], [19], [12].

======= Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. [3] [2] [4] [8] [5] [10] [14] [1] [7] [15] »»»> wildfire-test

2 Dataset

Our research utilizes the Wildfire Dataset, specifically Version 2 as published in the article *"The Wildfire Dataset: Enhancing Deep Learning-Based Forest Fire Detection with a Diverse Evolving Open-Source Dataset Focused on Data Representativeness and a Novel Multi-Task Learning Approach"* by El-Madafri et al. [9]. This dataset serves as a comprehensive benchmark for evaluating deep learning approaches to forest fire detection, addressing the critical need for reliable wildfire monitoring systems.

2.1 Data Sources and Composition

The Wildfire Dataset comprises 2,700 high-resolution RGB images collected from diverse sources including government databases (such as NASA and USGS), social media platforms like Flickr, and public domain image repositories such as Unsplash. These sources were selected to ensure geographic diversity across different forest ecosystems worldwide.

The dataset follows a two-class binary classification structure:

- **Fire**: Images showing evident fire-related phenomena
- **No Fire**: Forested areas without any signs of fire or smoke

The dataset dataset is provided in training (70%), validation (15%), and test sets (15%) with consistent class distribution across all subsets:

Table 1: Dataset partition distribution

Set	Total Images	Fire Images	No-Fire Images
Train	1,888	642	1,246
Validation	402	139	263
Test	410	109	301

This distribution maintained a roughly consistent class ratio of approximately 1 fire image for every 1.5-2 no-fire images across all partitions, preserving the original dataset's imbalance while allowing proper evaluation.

2.2 Dataset Characteristics

The Wildfire Dataset is characterized by substantial resolution variability, which presents both challenges and opportunities for deep learning models:

- Average Resolution: 4057×3155 pixels
- Minimum Resolution: 153×206 pixels
- Maximum Resolution: 19699×8974 pixels
- Standard Deviation (Width): 1867.47 pixels
- Standard Deviation (Height): 1388.60 pixels

This significant variability in image dimensions required careful preprocessing to ensure computational feasibility while maintaining the integrity of visual information.

2.3 Preprocessing Steps

Our preprocessing pipeline was designed to address the dataset's unique characteristics and enhance model performance:

2.3.1 Initial Processing

The images were organized into a directory structure with clear separation between training, validation, and test sets. Images are directly placed under "fire" or "nofire" subdirectories (as per Version 2 simplification), eliminating nested subdirectories to enhance accessibility and facilitate analysis.

2.3.2 Resolution Normalization

Given the substantial resolution range in the dataset, we conducted extensive preprocessing experiments to identify optimal image size. The final configuration selected during experimentation was 224×224 pixels after testing multiple resolutions (128×128 and 224×224 primarily). This decision was made empirically, as it provided the best results with the model we designed.

2.3.3 Data Augmentation

For the training phase, we experimented with the following augmentations:

- Rotation range: ± 15 degrees
- Horizontal shift range: up to 10% of width
- Vertical shift range: up to 10% of height
- Zoom range: up to 10%
- Horizontal flip (enabled)
- Brightness adjustment (range from 80%-120%)

This augmentation strategy aimed to increase model robustness while preventing overfitting, particularly important given the dataset's resolution variability. However, in our final iteration we did not end up using most of these. [TODO: which did we actually use?]

2.3.4 Data Normalization

All images were normalized with rescaling by a factor of $1/255.0$ to ensure consistent input range for the neural network. This preprocessing step standardizes pixel values across the entire image set and improves model convergence.

The dataset is licensed under CC BY 4.0, which allows use for our purposes with appropriate attribution to the original authors and publication.

2.4 Dataset Evolution and Quality Assurance

This version of the Wildfire Dataset represents a carefully curated collection designed for machine learning applications in forest fire detection. The researchers specifically focused on data representativeness by including:

- Images from diverse geographic locations across different continents
- A wide range of environmental conditions (e.g., time of day, weather conditions)
- Different types of forest ecosystems and vegetation

- Various wildfire stages, from early ignition to fully developed flames

3 Features and Inputs

The input to all models consists of raw RGB image pixels, treated as the feature vector for each example. Because CNNs require inputs of uniform dimensionality, all images were resized to a fixed resolution prior to training. Multiple resolutions were tested: 128×128, 224×224, 299×299, and 1000×1000. These examined how increasing the number of pixels (and thus the feature dimensionality) affects model performance. The specific numbers for pixel sizes were chosen based on commonly discussed and used pixel sizes for CNN's found in the various cited online sources [17], [18]. However, resolutions above 224×224 caused out-of-memory (OOM) failures on our RAM due to the rapid growth in tensor sizes and batch memory requirements. We tried running the higher resolution experiments on Google Cloud as well but due to limitations on our billing account, we were unable to allocate sufficient GPU resources to handle the larger resolutions as well. As a result, larger resolutions were excluded from final experiments.

During preprocessing, images were normalized to the [0, 1] range to stabilize training and improve convergence. A manual inspection of the dataset revealed a small number of incorrectly labeled images, which introduced noise into the feature space; this observation further motivated the use of robust architectures and augmentations that help models generalize under imperfect labels. To enhance the diversity of the training samples and make the models invariant to orientation, position, and scale, we applied several image augmentations: random rotations, width and height shifts, zooming, and horizontal flips. The idea to vary the pixel size and augmentations was inspired from an online project tutorial which developed a Malaria Detection model using TensorFlow's malaria dataset [20], [6]. The scope and scale of the Malaria Detection project were similar to ours, and we adapted their augmentation strategies to our wildfire dataset. These transformations effectively vary the spatial arrangement of pixel features while preserving semantic content, allowing the models to learn more resilient representations. Overall, no handcrafted feature engineering was performed; instead, the CNN architectures learn hierarchical feature representations directly from pixel data. By varying both image resolution and augmentation strategies, we were able to study how feature dimensionality and data variability affect wildfire detection performance across different models.

4 Implementation

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. [1]

5 Evaluation

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. [1]

6 Progress

The original plan was to evaluate the current wildfire detection model and identify areas for improvement. We initially focused on three key aspects: efficiency, accuracy, and mobility. Accuracy was the most straightforward area to target, since increasing correct predictions—especially reducing false negatives—is critical for wildfire detection. Efficiency mattered because a faster model reduces resource usage and shortens the time needed to confirm whether a wildfire has begun. Mobility was also important, as we wanted to explore whether the model could eventually be deployed on more portable and cost-effective hardware rather than relying on expensive systems.

To follow through on this plan, we implemented and compared three different models—ResNet18, EfficientNet, and MobileNet—using the same dataset. Each model was chosen because it excels in one of the improvement categories. Based on feedback from my previous progress report, We also incorporated additional evaluation metrics, such as ROC/AUC, to better capture the strengths and weaknesses of each approach.

However, the plan shifted slightly during the process. After running the models, it became clear that more detailed error analysis was needed, especially to address the high rate of false negatives. Because of this, the direction of the project moved from general improvement across multiple categories to a more targeted focus on understanding and reducing misclassifications.

7 Error Analysis

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. [1]

Data Description: missing class distribution details

Implementation: missing architectural details

Evaluation: limited error analysis, missing ROC/AUC

Feedback and Next Step: no strategy for addressing false negatives.

Sapna: Overall, good work!

Team Contributions

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. [1]

References

[1] National Aeronautics and Space Administration. 2025. Wildfires and climate change. Web page. Available at https://science.nasa.gov/earth/explore/wildfires-and-climate-change/. Accessed: 2025-11-09.

[2] Alfred V. Aho and Jeffrey D. Ullman. 1972. *The Theory of Parsing, Translation and Compiling*, volume 1. Prentice-Hall, Englewood Cliffs, NJ.

[3] Arize AI. 2023. Binary cross entropy: Where to use log loss in model monitoring. https://arize.com/blog-course/binary-cross-entropy-log-loss/. Published January 30 2023.

[4] American Psychological Association. 1983. *Publications Manual*. American Psychological Association, Washington, DC.

[5] Galen Andrew and Jianfeng Gao. 2007. Scalable training of L1-regularized log-linear models. In *Proceedings of the 24th International Conference on Machine Learning*, pages 33–40.

[6] Bnsreenu. 2023. Malaria binary classification using tensorflow lite. https://github.com/bnsreenu/python_for_microscopists/tree/master/237_tflite_using_malaria_binary_classification. GitHub repository, accessed on 2025-11-10.

[7] Health Canada. 2024. Human health effects of wildfire smoke — report summary. Web document. Available at https://www.canada.ca/en/services/health/healthy-living/environment/air-quality/wildfire-smoke/human-health-effects-report-summary.html. Accessed: 2025-11-09.

[8] Ashok K. Chandra, Dexter C. Kozen, and Larry J. Stockmeyer. 1981. Alternation. *Journal of the Association for Computing Machinery*, 28(1):114–133.

[9] I. El-Madafri, M. Peña, and N. Olmedo-Torre. 2023. The wildfire dataset – enhancing deep learning-based forest fire detection with a diverse evolving open-source dataset. https://www.kaggle.com/datasets/elmadafri/the-wildfire-dataset?resource=download. Kaggle dataset, accessed on 2025-11-10.

[10] Dan Gusfield. 1997. *Algorithms on Strings, Trees and Sequences*. Cambridge University Press, Cambridge, UK.

[11] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.

[12] Andrew G Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.

[13] Rose Morrison. 2022. The impact of wildfires on biodiversity and the environment.

[14] Mohammad Sadegh Rasooli and Joel R. Tetreault. 2015. Yara parser: A fast and accurate dependency parser. *Computing Research Repository*, arXiv:1503.06733. Version 2.

[15] Laxita Soontha and Mohammad Younus Bhat. 2026. Global firestorm: Igniting insights on environmental and socio-economic impacts for future research. *Environmental Development*, 57. Accessed: 2025-11-10.

[16] Dario Spiller, Andrea Carbone, Stefania Amici, Kathiravan Thangavel, Roberto Sabatini, and Giovanni Laneve. 2023. Wildfire detection using convolutional neural networks and prisma hyperspectral imagery: A spatial-spectral analysis. *Remote Sensing*, 15(19):4855.

[17] Stack Overflow user Ivan Shelonik. 2018. What should be the dimension of image in convolutional neural network? https://stackoverflow.com/questions/48954724/what-should-be-the-dimension-of-image-in-convolutional-neural-network. Accessed: 2025-12-03.

[18] Stack Overflow user user10024395. 2017. Is there any particular reason why people pick 224×224 image size for imagenet experiments? https://stackoverflow.com/questions/43434418/is-there-any-particular-reason-why-people-pick-224x224-image-size-for-imagenet-e. Accessed: 2025-12-03.

[19] Mingxing Tan and Quoc V Le. 2019. Efficientnet: Rethinking model scaling for convolutional neural networks. *arXiv preprint arXiv:1905.11946*.

[20] TensorFlow Datasets. 2023. Malaria dataset. https://www.tensorflow.org/datasets/catalog/malaria. Contains 27,558 thin blood-smear cell images with parasitized/uninfected labels.

[21] Eleni Tsalera, Andreas Papadakis, Ioannis Voyiatzis, and Maria Samarakou. 2023. Cnn-based, contextualized, real-time fire detection in computational resource-constrained environments. *Energy Reports*, 9:247–257.

[22] Berk Özel, Muhammad S. Alam, and Muhammad U. Khan. 2024. Review of modern forest fire detection techniques: Innovations in image processing and deep learning. *Information*, 15(9):538.