Report:Wildfire Prediction using Neural Networks

Introduction: Wildfires are one of the most destructive natural disasters,they possess a severe threat to the ecosystem,climate,habitat and communities residing near the area.
Accurate wildfire predictions could help emergency services respond effectively ,reducing damage and saving lifes.
This project aims to predict the location of these wildfires using deep learning techniques that would be applied on a Satellite Images dataset of the Canadian Surface.

Dataset:https://www.kaggle.com/datasets/abdelghaniaaba/wildfire-prediction-dataset/
This dataset contains satellite images (350x350px) in 2 classes :

Wildfire : 22710 images
No wildfire : 20140 images
The data was divided into train, test and validation with these percentages :

Train : ~70% (30250 images)
Test : ~15%   (6300 images)
Validation : ~15% (6300 images)

The 2 main objectives of this project are to
1. Develop a custom Convolutional Neural Network(CNN)
2. Develop a pretrained model using transfer learning(ResNet50)

Literature Review:
1. Forest Fire prediction using Machine Learning Methods:
   This paper presents a comparative study of four popular ML methods-decision tree, random forest, k-nearest neighbors (KNN), and support vector machine (SVM)-for forest fire detection, in terms of accuracy, precision, recall, and F1 score. The study also compared the test duration of each method.The experimental results show that the decision tree outperforms the other three algorithms, achieving an accuracy of 97.95%, a precision of 100%, a recall of 97.05%, and an F1 score of 98.5%.

2. Deep learning models for enhanced forest-fire prediction at Mount Kilimanjaro, Tanzania:

   The study evaluates advanced Deep Learning (DL) models for FF prediction by integrating spatiotemporal vegetation indices, environmental data, and human activity indicators. Specifically, Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNNs), and Convolutional Long Short-Term Memory (ConvLSTM) models were employed to analyze Sentinel-2 satellite imagery and weather data.The ConvLSTM model engineered to capture intricate spatial and temporal relationships delivered superior performance, achieving an AUROC of 0.9785 and Accuracy 98.08%, surpassing the LSTM and CNN models.

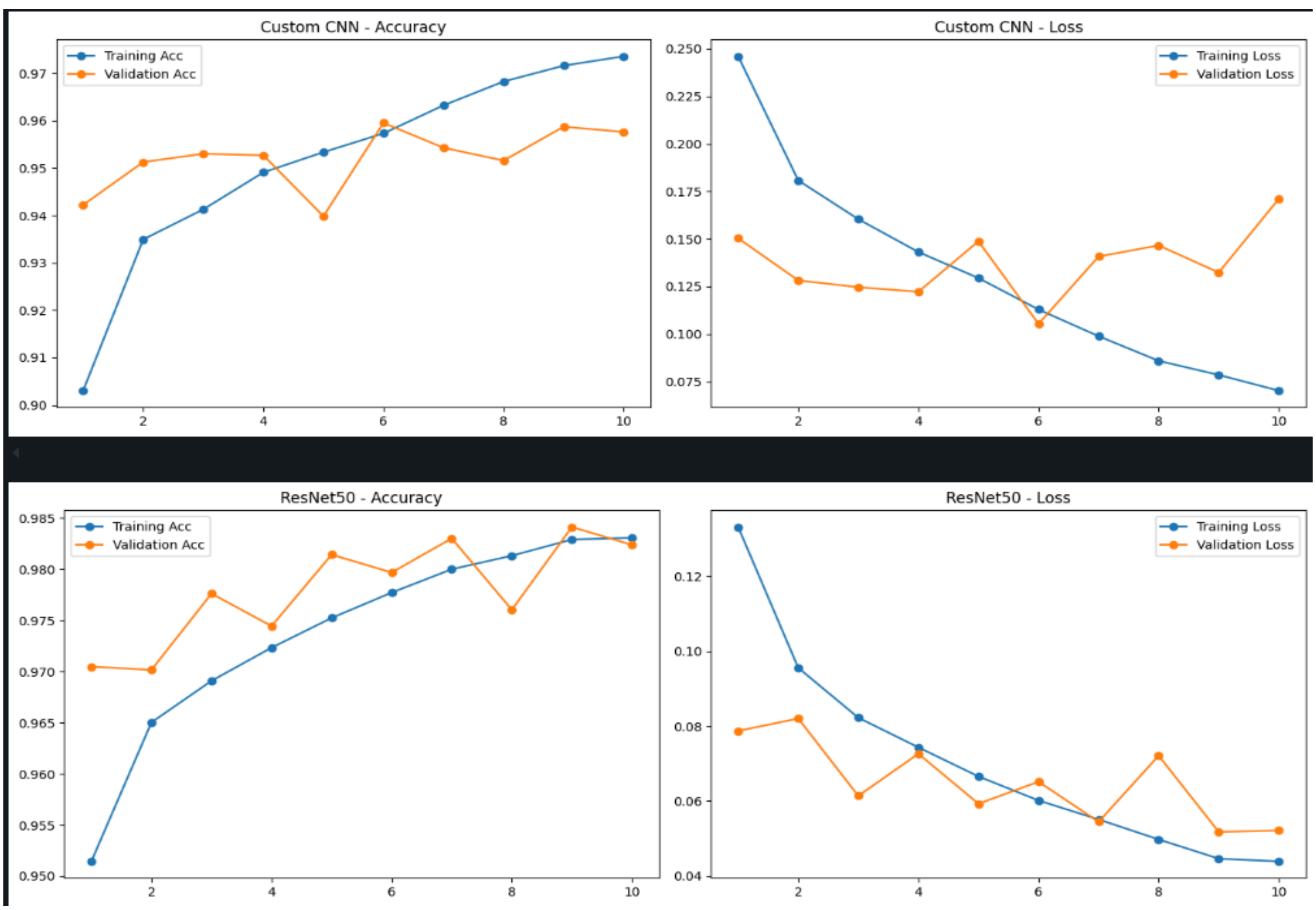
Methodology

1. Data Cleaning and Preprocessing:
    1. The corrupted images are deleted and if some corrupted images still remain it is taken care that these images are partially loaded using Pillow(PIL) library which ensures that model does not gets crashed during training.
    2. A Pandas dataframe is created for each of the train,test & validate folders **Columns**: **Image**(gives the path of the image file) and **Label**(nowildfire/wildfire image).
    3. All the images are **Normalized** and **scaled to [0,1] range.**

2. Custom CNN Architecture:
    1. A sequential model is developed with input size **128x128 rgb image**.
    2. Model consists of 2 convolutional layers with 1st one with **32 filters** and 2nd with **64 filters** where size of each **filter is (3x3)** both the layers contain a **ReLU activation** function .
    3. Each of the convolutional layers is followed with a **(2x2) max pooling layer.**
    4. The feature maps are converted into **1D vector** using the .**Flatten**
    5. A **Dense layer** with **64 neurons** and **ReLU activation** is used to learn higher-level representations
    6. **50%** of the neurons are stopped to reduce overfitting by using the **Dropout Layer**
    7. The final layer is a Dense layer with **1 neuron** and a **sigmoid activation**, which outputs a probability for **binary classification**.
    8. The model is compiled with the **Adam optimizer**, **binary cross-entropy** loss (as it is a binary classification problem), and **accuracy** as the performance metric, it is trained for 10 epochs.

3. ResNet50 Architecture:
    1. The model accepts **224x244 rgb images**(by default)
    2. The base model is **ResNet50**(without the classification layer) this extracts deep hierarchical features from the input images
    3. A **Global Average Pooling layer** is used to reduce spatial dimension of feature maps.
    4. Then A **Dense layer with 128 neurons and ReLU activation** captures high-level information.
    5. Like CNN **50%** of the neurons are stopped to reduce overfitting by using the **Dropout Layer.**
    6. Finally, a **Dense output layer with 1 neuron and sigmoid activation** produces a binary prediction- wildfire or no wildfire.
    7. The model is compiled with the **Adam optimizer**, **binary cross-entropy** loss (as it is a binary classification problem), and **accuracy** as the performance metric, it is trained for 10 epochs.

Results:

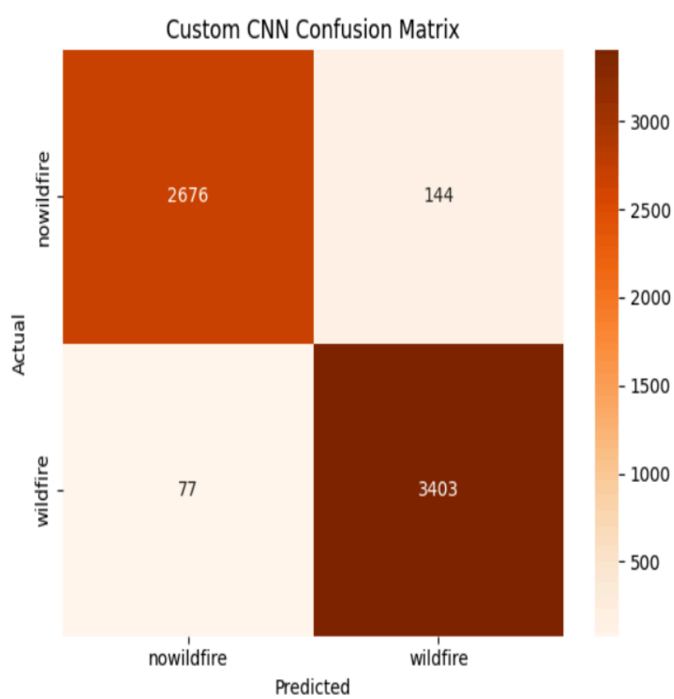


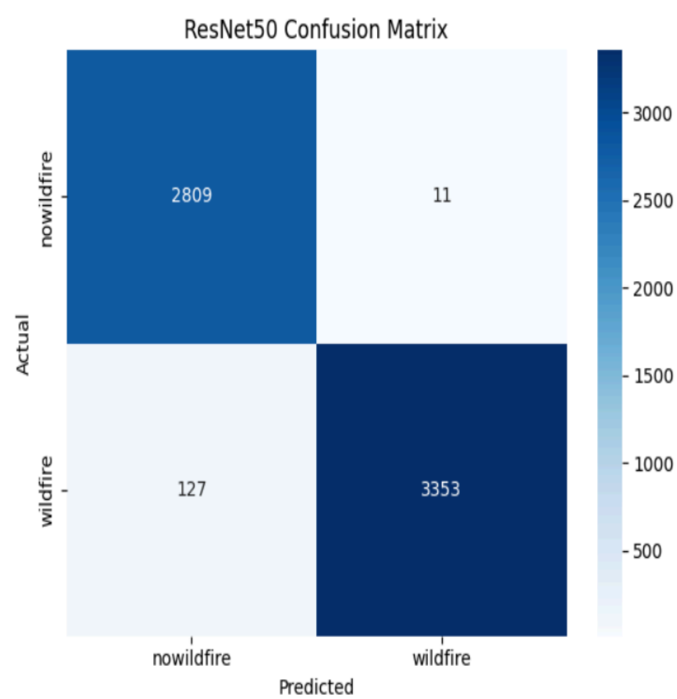
- The Custom CNN showed good improvement during training, but after around epoch 6, the validation accuracy started to fluctuate and the loss didn't drop much—so there's a bit of overfitting going on.
- On the other hand, ResNet50 gave consistently high performance on both training and validation sets, which means it generalized better and learned more stable features overall**.**

Metrics Over Testing Dataset

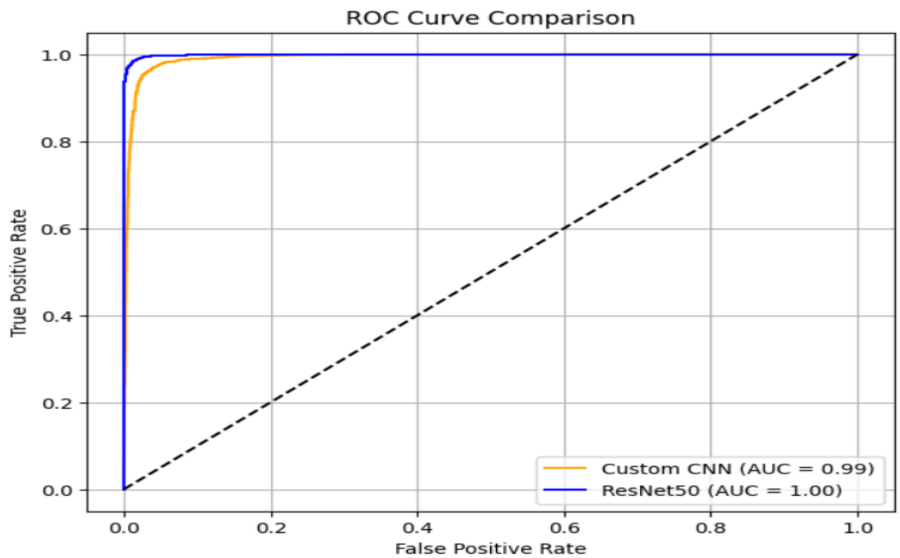
	Accuracy	Precision	Recall	F1-Score
Custom CNN	0.964921	0.959402	0.977874	0.968550
ResNet50	0.978095	0.996730	0.963506	0.979836

Confusion Matrix





- The custom CNN holds up pretty well, but it still slips up on 144 'nowildfire' and 77 'wildfire' cases, so there is definitely room to fine-tune its precision and sensitivity.
- ResNet50, absolutely nails it with just 11 'nowildfire' and 127 'wildfire' misclassifications, making it more reliable and accurate overall.



Both models did really well, but ResNet50 stood out with an AUC of 1.00 — showing how accurately it can separate wildfire from non-wildfire images.

Conclusion:
This project compared a custom CNN model against a ResNet50-based transfer learning model for wildfire image classification.A Custom CNN(Accuracy 96.49%) model performed very well as compared to a pre-trained ResNet50(Accuracy 97.80%) model.

References:

- Sharma, N., Jain, V., & Mishra, A. (2018). An analysis of convolutional neural networks for image classification. *Procedia computer science*, *132*, 377-384.
- Mambile, C., Kaijage, S., & Leo, J. (2024). Deep Learning Models for Enhanced Forest-Fire Prediction at Mount Kilimanjaro, Tanzania: Integrating Satellite Images, Weather Data and Human Activities. *Natural Hazards Research*.
- Preeti, T., Kanakaraddi, S., Beelagi, A., Malagi, S., & Sudi, A. (2021, June). Forest fire prediction using machine learning techniques. In *2021 International Conference on Intelligent Technologies (CONIT)* (pp. 1-6). IEEE.
- Aarich, M., Rouijel, A., & Amine, A. (2024). Deep Learning Approaches for Forest Fires Detection and Prediction using satellite Images. *Procedia Computer Science*, *251*, 758-763.