Wildfire Dataset

Danish Bhatkar
Clemson University
School of Computing
Clemson, SC, USA
dbhatka@clemson.edu

I. Introduction

The Wildfire Dataset project involves applying multiple image processing techniques to multiple datasets and comparing the results. This study involves only the classification tasks. The ResNet model is used for training on multiple datasets with the same structure, optimizer, loss function, learning rate, etc., and then test the trained model on the remaining datasets.

The idea is to analyze which dataset is the most optimal and can generalize information across all other datasets.

II. Datasets used for Training

A. FLAME

The training dataset consists of 39,375 frames resized to 254x254 for the "Fire-vs-NoFire" image classification problem (Training/Validation dataset). The size of this repository is 1.3 GB and the format is JPEG. It is divided into a 3:1 ratio for training and validation purposes. The testing dataset consists of 8,617 frames resized to 254x254 for the "Fire-vs-NoFire" image classification problem (Test dataset). The size of this repository is 301 MB and the format is JPEG.

B. FLAME2

The dataset consists of 53,541 side-by-side RGB/IR frame pairs. The frames have then been scaled down to 254p x 254p and cropped such that both the RGB/IR frame possess a similar FOV and perspective. There are 53,451 RGB frames. The frame pairs are labeled with a "Fire/NoFire" label and a "Smoke/NoSmoke" label. For training purposes, only the RGB images have been considered and split into training, validation and evaluation sets in the ratio 3:1:1.

C. DeepFire

All images in the dataset are 3-channeled with resolution of 250 × 250. The images were retrieved by searching various search terms in multiple search engines. Afterwards, these images are thoroughly investigated to crop and remove the inappropriate components such as people, fire-extinguishing machinery etc. to ensure that each image only contains the relevant fire region. The dataset is designed for binary problem of Fire and No-Fire detection in the forests landscape. It is a balanced dataset consisting of 1900 images in total, where 950 images belong to each class. The dataset is divided into 80:20 for training and testing purposes in the proposed study.

D. Wildfire (Kaggle)

All images in the dataset are 3-channeled with resolution of 250 × 250. The dataset is designed for binary problem of Fire and No-Fire detection in the forests landscape. It has 1832 training images and 68 test images.

E. D-Fire

D-Fire is an image dataset of fire and smoke occurrences designed for machine learning and object detection algorithms with more than 21,000 images, categorized into 'Only fire', 'Only smoke', 'Fire and Smoke', and 'None'. All images were annotated according to the YOLO format (normalized coordinates between 0 and 1).

F. FF-Det

All images in the dataset are 3-channeled with spatial resolution of 250 × 250. Inappropriate components such as people, fire-extinguishing machinery etc, have been removed. Each image only contains the relevant fire region. The dataset is designed for binary problem of fire or no-fire detection in the forests landscape. It is a balanced dataset consisting of 1900 images in total, where 950 images belong to each class. The dataset is divided into 80:20 for training and testing purposes in our study.

G. FLAME3

FLAME 3 is the third dataset in the FLAME series of aerial UAV-collected side-by-side multi- spectral wildlands fire imagery (see FLAME 1 and FLAME 2). This set is a single -burn subset of the Larger FLAME 3 dataset focusing specifically on Computer Vision tasks such as fire detection and segmentation. Included are 622 image quartets labeled Fire and 116 image quartets labeled No Fire. The No Fire images are of the surrounding forestry of the prescribed burn plot. Each image quartet is composed of four images - a raw RGB image, a raw thermal image, a corrected FOV RGB image, and a thermal TIFF.

III. Training Methodology

A. Image Transformations

The following transformations were applied for making the dataset compatible with ResNet-50's input structure:-

- Resized images to 224x224 pixels.
- Standard normalization using ResNet-50's recommended mean and standard deviation values to match the pre-trained model weights.

B. Model used for training

1) Architecture: A pre-trained ResNet-50 model was used and modified by replacing the output layer (fully connected layer) with a new layer for two output classes (Fire and No Fire).

2) Training Parameters:

- Loss Function: Cross-Entropy Loss
- Optimizer: Adam optimizer with a learning rate of 0.001
- Batch Size: 16
- Epochs: 20

IV. Considerations while training and testing the models

1) FLAME2 – 20% of RGB images used for testing
2) Bowfire – Images from 'dataset' directory used for testing
3) FF-Det – No images in training directory. Used test directory for training the model.
4) FiSmo – Used images from Flick-Fire directory for testing.
5) FLAME3 – Images from 'RGB/Corrected FOV' directories were used for both fire and no fire classes for training and testing purposes

Number of images used from each dataset for testing:-

- FLAME - 8617
- FLAME2 - 10690 (20% of total images)
- DeepFire - 380
- DataCluster - 100
- FF-Det - 380
- Fire - 1000
- Wildfire - 68
- BowFire - 226
- D-Fire - 4306
- FiSmo - 1000
- FLAME3 - 738

V. Results

A. FLAME

Training Loss = 0.0030 Training Accuracy = 99.87%

Dataset	Model Accuracy	Precision	Recall	F1
FLAME2	74.70%	0.75	1.00	0.86
DeepFire	98.42%	1.00	0.98	0.99
DataCluster	93.94%	1.00	0.94	0.97
FF-Det	48.66%	0.49	0.94	0.65
Fire	23.95%	0.23	0.92	0.37
Wildfire	66.18%	0.68	0.96	0.79
BowFire	92.48%	1.00	0.92	0.96
D-Fire	53.74%	0.54	0.95	0.69
Fismo	95.90%	1.00	0.96	0.98
FLAME3	84.28%	0.84	1.00	0.91

B. FLAME2

Training Loss = 0.0021 Training Accuracy = 99.95%

Dataset	Model Accuracy	Precision	Recall	F1
FLAME	40.39%	0.16	0.40	0.23
DeepFire	48.95%	1.00	0.49	0.66
DataCluster	52.53%	1.00	0.53	0.69
FF-Det	51.32%	0.51	0.46	0.49
Fire	23.25%	0.02	0.05	0.03
Wildfire	44.12%	0.61	0.48	0.54
BowFire	49.56%	1.00	0.50	0.66
D-Fire	50.81%	0.55	0.41	0.47
Fismo	58.10%	1.00	0.58	0.74
FLAME3	30.49%	0.69	0.31	0.43

C. DeepFire

Training Loss = 0.0929 Training Accuracy = 96.19%

Dataset	Model Accuracy	Precision	Recall	F1
FLAME	45.36%	0.67	0.45	0.35
FLAME2	68.21%	0.77	0.82	0.79
DataCluster	35.35%	1.00	0.35	0.52
FF-Det	96.84%	0.97	0.96	0.97
Fire	95.81%	0.87	0.98	0.92
Wildfire	98.53%	0.98	1.00	0.99
BowFire	25.22%	1.00	0.25	0.40
D-Fire	40.90%	0.46	0.68	0.55
Fismo	4.10%	1.00	0.04	0.08
FLAME3	84.55%	0.88	0.95	0.91

D. Wildfire (Kaggle)

Training Loss = 0.0152 Training Accuracy = 99.56%

Dataset	Model Accuracy	Precision	Recall	F1
FLAME	40.39%	0.16	0.40	0.23
FLAME2	67.50%	0.73	0.90	0.81
DeepFire	48.42%	1.00	0.48	0.65
DataCluster	35.35%	1.00	0.35	0.52
FF-Det	97.89%	0.97	0.99	0.98
Fire	96.21%	0.92	0.93	0.92
BowFire	30.97%	1.00	0.31	0.47
D-Fire	41.08%	0.46	0.68	0.55
Fismo	5.50%	1.00	0.06	0.10
FLAME3	83.20%	0.84	0.99	0.91

E. D-Fire

Training Loss = 0.0565 Training Accuracy = 98.03%

Dataset	Model Accuracy	Precision	Recall	F1
FLAME	40.15%	0.33	0.40	0.36
FLAME2	69.16%	0.78	0.83	0.80
DeepFire	59.47%	1.00	0.59	0.75
DataCluster	82.83%	1.00	0.83	0.91
FF-Det	29.21%	0.35	0.47	0.40
Fire	11.38%	0.08	0.23	0.11
Wildfire	32.35%	0.50	0.41	0.45
BowFire	71.24%	1.00	0.71	0.83
Fismo	90.10%	1.00	0.90	0.95
FLAME3	84.01%	0.87	0.95	0.91

F. FF-Det

Training Loss = 0.0192 Training Accuracy = 99.30%

Dataset	Model Accuracy	Precision	Recall	F1
FLAME	58.96%	0.43	0.59	0.45
FLAME2	42.18%	0.63	0.56	0.59
DeepFire	45.53%	1.00	0.46	0.63
DataCluster	23.23%	1.00	0.23	0.38
Fire	94.41%	0.89	0.88	0.88
Wildfire	100.00%	1.00	1.00	1.00
BowFire	27.43%	1.00	0.27	0.43
D-Fire	39.46%	0.45	0.56	0.50
Fismo	3.00%	1.00	0.03	0.06
FLAME3	63.69%	0.80	0.75	0.78

G. FLAME3

Training Loss = 0.0409 Training Accuracy = 98.19%

Dataset	Model Accuracy	Precision	Recall	F1
FLAME	57.33%	0.51	0.52	0.57
FLAME2	69.80%	0.80	0.79	0.80
DeepFire	42.63%	1.00	0.43	0.60
DataCluster	29.29%	1.00	0.29	0.45
FF-Det	65.79%	0.66	0.64	0.65
Fire	61.48%	0.24	0.27	0.25
Wildfire	72.06%	0.80	0.78	0.79
BowFire	56.64%	1.00	0.57	0.72
D-Fire	47.75%	0.51	0.65	0.57
Fismo	19.40%	1.00	0.19	0.33

VI. Summary of the Generalization Results

Section V shows the generalization pattern of the 7 datasets on which the ResNet model is trained for the fire and no fire classification task. Each trained model is tested on all the datasets but the training dataset itself.

To summarize all the results, the average testing accuracy is calculated for each training dataset and recorded in table I. For example, all the testing accuracies obtained for the FLAME dataset are averaged out and displayed against the FLAME dataset in table I. The same is repeated for the other six datasets, too.

Apart from the test accuracy, the F1 score is considered too because it is a blend of Precision and Recall, and provides a good idea about the model performance. Precision and Recall are not taken into account since they focus on one aspect of the model performance and could be misleading.

Dataset	Average Testing Accuracy	Average F1 Score
FLAME	73.23%	0.82
DeepFire	59.49%	0.65
D-Fire	56.99%	0.65
Wildfire	54.65%	0.61
FLAME3	52.22%	0.57
FF-Det	49.79%	0.57
FLAME2	44.95%	0.50

TABLE I
Average Testing Accuracies and F1 Scores

VII. Conclusion

Table I is ordered by the testing accuracy values with the highest value on the top.

Based on the values, the ResNet model trained on the FLAME dataset has the highest average testing accuracy (73.23%) and the highest F1 score (0.82) when tested on other models. It can be concluded that FLAME is the best dataset (among the 7 datasets used for training) in terms of generalization results.

VIII. Important Links

1) GitHub
2) Datasets