Classification of fire and non-fire aerial images Using Transfer Learning on FLAMES Data set

JhansiRani Choutapalem

Department of Mathematical Sciences,University of Essex
Registration Number : 2004458
Email:jc20168@essex.ac.uk

[1] ***Abstract*—Forest fires pose greater threat to environment and ecosystem. Early detection and prevention of fires before they become uncontrollable is vital to reduce the damages caused by fire. Several studies have shown that Deep learning networks can accurately predict fire in images. This highlights the use of deep neural networks such as Convolution Neural Networks(CNN) as fire detection systems, to identify the presence of fire in images and videos captured in real time. Traditional CNN's however, takes long time to train and require large amount of data to make accurate predictions. This paper proposes use of transfer learning technique to build fire detection models.The data set used for this study is FLAMEs data set, which contains raw video footage recorded by drones during burning of detritus piles in Arizona forest, USA. The video is converted into frames and used for binary image classification ("Fire" vs "No-fire"). To evaluate this research question, two models were built, the first model is a traditional deep CNN, which was able to achieve an accuracy of 66% on test set. The second model is built by using and fine-tuning existing Xception net and it has achieved an accuracy of 82.6% on test data. The experimental results shows that transfer learning achieves higher accuracy compared to traditional CNN for image classification.**

***Index Terms*—Forest fires; deep CNN; transfer learning; Xception net**

I. Introduction

Fire is a major natural disaster that can have severe impact on ecosystem. Fires in the forest causes loss of biodiversity destroying vegetation and natural habitat of animals. In addition, smoke from fire causes respiratory and eye illness to the operational forces and residents of local area.

In USA, according to National Inter agency Fire Centre(NIFC), from years 2010 to 2019, there were 50,417 fires which destroyed 4,664,364 acres of forest land, accounting for more than $18 billion suppression costs[1]. In addition, the management of fires is expensive and often pose risk to lives of forces involved in fire control[2]. Recent research suggests that due to climate change and global warming the world may witness a greater number of wildfires in the coming years[3].

This highlights the importance of fire detection in its early stages, as early detection[4,5] plays crucial role in reducing the damage caused by the fire to humans and nature. Continuous monitoring of the forest land is also essential in order to identify the fire location.

Traditional fire detection systems involve usage of equipment such as thermal or smoke detectors installed randomly in the forest, this process is laborious, expensive and infeasible considering massive size of some forests. In contrast, use of latest technology such as unmanned aerial vehicles(UAVs), have gained a lot of popularity in recent times because of its beyond human capabilities in surveillance over massive forest land[6,7,8,9,10].

Fire detection system with real time image capturing, predicting the fire from captured images and alerting the fire management system is considered a promising approach to detect fire anywhere in the forest and prevent it from causing heavy damages. Deep learning networks such as Convolution Neural Networks have been used widely for image classification and object detection. However, there are not many deep learning models that are available for fire detection due to scarcity of data.

For dealing with data scarcity problem, transfer learning method is used. Transfer learning enables use of knowledge gained on one problem to solve a similar but different problem.

In this paper, Xception[18] model pretrained on Image Net data set is used for transfer learning and the results of transfer learning is compared to that of traditional CNN.

II. Literature Review

Traditional image processing methods involved comparing RGB values[11] of different objects such as fire in images and videos. However, these methods are not fully reliable owing to their disability to not distinguish between various objects as they solely rely on colour intensity. Convolution neural networks are specifically designed for image detection and classification tasks. CNN's can identify and classify different objects, thereby overcome the problems with conventional approaches.

The following work demonstrates some of the published work done in this domain.

In this paper[12], an architecture similar to LeNet5 is implemented by authors for detecting and classifying smoke and fire in video footage. The architecture has achieved an accuracy of 97.9% on test data.

A model using AlexNet architecture to identify smoke in images has developed in paper[13]. The model was trained on large number of images containing smoke and no-smoke and

[1]Department of Mathematical Sciences, University of Essex, CO4 3SQ, Colchester, United Kingdom (e-mail:jc20168@essex.ac.uk).

has achieved an accuracy of 96.88% with low false positive rate. In[14], the authors artificially created images and videos using blender and used AlexNet model for training this data. However, the model was tested on real data. The model has achieved reasonable accuracy, but the false positive rate is high.

In this paper[15], the authors have created a concise version of Xception model. This model has 3 layers, an input layer, hidden layers and output layer. The input layer depends on image size and number of channels, re scaling is performed. The hidden layers depend on depth-wise separable convolution layers and a short cut is added between convolution blocks like ResNet model. The output layers contain a dense layer followed by a classification layer with sigmoid as activation function. The paper has achieved an accuracy of 76% on test data.

The authors of this paper [16], used transfer learning technique to detect fire in images. The data set is hand curated by authors by taking fire and non-fire images from various sources like Image Net, Baidu, Google, and capturing fire frames from YouTube videos. Data augmentation is performed to increase the amount of data available for training. The authors chose Xception and InceptionV3 for transfer learning. They added a dropout layer and dense layer on top of pretrained model, then they freezed the pretrained model and only trained the dense layers. After that, they fine tuned this model by unfreezing the last two convolution blocks and trained it again on the whole data set. The authors reported 96% accuracy on test data after fine-tuning.

In [17], authors performed transfer learning using InceptionV3 and MobileNetv2 models to detect fire in surveillance videos. The data is curated by authors by taking fire videos from internet and own recordings. Their experiment involved comparing full model training versus transfer learning followed by fine tuning. Results showed that transfer learning is better than full training in both the models they tried.

From the literature review, it is evident that Deep CNN's showed promising results for the object detection and image classification tasks compared to traditional approaches. Traditional methods are computationally inexpensive, but they suffer either from low accuracy or high false positive rate problem. Even though Deep CNN's are computationally expensive, when trained on high volume data, they achieve high accuracy. Transfer learning over comes the problem of data scarcity and using limited data, transfer learning models can achieve high accuracy, which otherwise is not possible with conventional methods.

The work done in this paper is inspired from previous researchers work[15,16,17] done in the field of fire detection in images and videos using transfer learning and this proposes to use Xception net for transfer learning and fine tuning it to classify fire and non-fire images in aerial images from FLAMES data set and also, compares the transfer learning model with traditional deep CNN.

III. Methodology

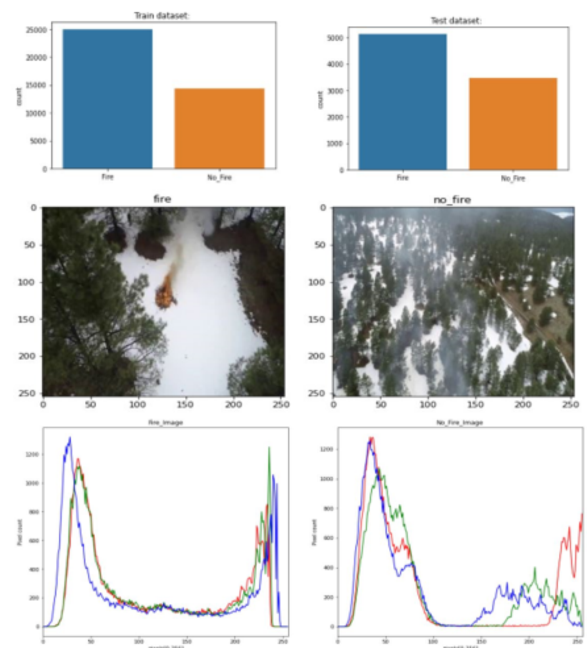
The seventh repository in FLAMES data set[15] is used for fire vs non-fire image classification. It has 39,375 frames that are labelled for training phase and 8,617 frames are labelled for testing. The size of training repository is 1.3GB and size of test data is 301MB. Each image has dimensions of 254×254, with a depth of 3. The table below shows the details of data sets:

Data set	Image Format	Number of Images	Size
Training/validation	JPEG	39375	1.3GB
Test	JPEG	8617	301MB

The training data set has 25,018 fire images and 14,357 non-fire images. 63.5% of training data contains fire images, and 36.5% of data contains non-fire images, this is a quite imbalanced data set.

The figure below shows the basic data analysis performed on this data:

Fig. 1. Data Analysis



For both CNN model and transfer learning with Xception, ImageDataGenerator class from Keras library is used to read the data and to perform preprocessing. Using ImageDataGenerator, data augmentation pipeline is created for training data. The augmentation involved width shift, height shift, horizontal flip, vertical flip and re scale. Also, 20% of training data is split into validation data to validate the model at the end of every epoch.Batch size of 64 is used for training and validation. For testing batch size of 1 is used.

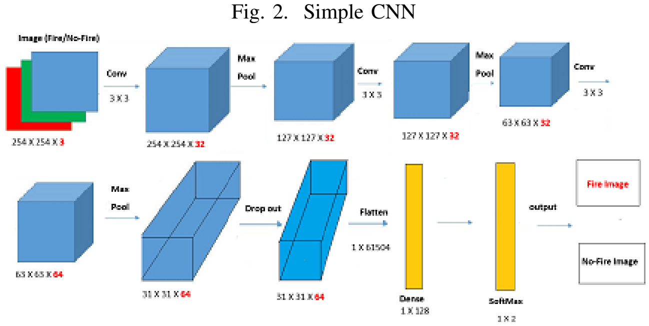
Using Keras callbacks Early stopping is defined to avoid model from over fitting. Additionally,Model Checkpoint is created to stop training when the validation loss is not improving after a given threshold(10 epochs). ReduceLROnPlateau callback is used to reduce the learning rate if the loss is not decreasing after a given threshold(10 epochs).

Adam was chosen as optimizer, Binary cross entropy as loss function and accuracy as performance metric. Since this is a binary classification problem, the activation function for last layer in both models is Sigmoid function.

A. *Baseline model(CNN):*

A CNN model is constructed with 3 Convolution layers, each followed by an activation layer with relu as activation function and a Max pooling layer. A Flatten layer is added after third convolution block, followed by a Dense layer. Dropout layer is added to avoid over fitting, it is followed by a fully connected dense layer with sigmoid activation function.

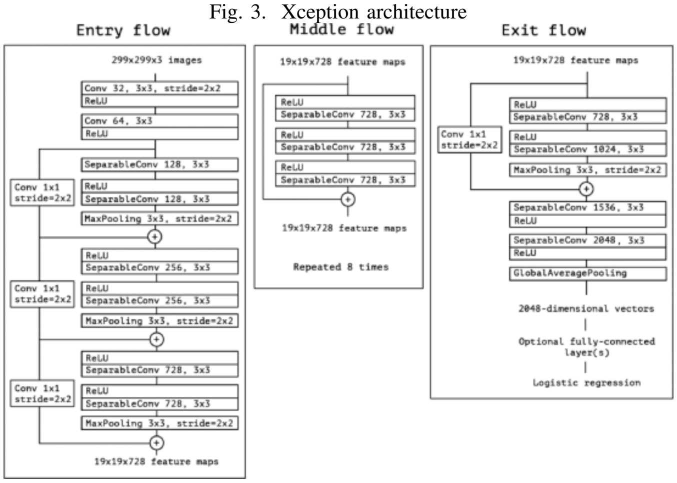
Fig. 2. Simple CNN



B. *Xception:*

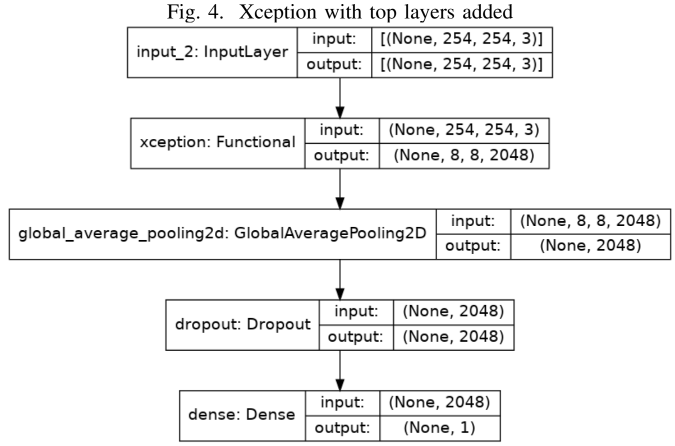
Xception stands for Extreme Inception, this architecture resulted from replacing standard Inception modules with depth wise separable convolutions. Xception architecture consists of linear stack of depth wise separable convolution with residual connections. In total it has 36 convolution layers that are structures into 14 modules, except for the first and last modules, every other module has residual connections around them. Xception is a lightweight model with 88MB size and has won Image Net challenge with 79% accuracy against Inception3, VGG16 and Resnet.

Fig. 3. Xception architecture



1) **Transfer learning**: For the transfer learning the following steps are followed:
   - First Base model(Xception) is instantiated with includetop = False to remove the top classification layer, and pretrained weights are loaded into it.
   - All the layers of the base model are frozen by setting trainable=False.
   - A new model is created by adding a Global Average Pooling layer, a Dropout layer and a Dense layer on top of output of base model.
   - This new model is trained on new data set(FLAMES). This model is trained for 20 epochs, however due to early stopping the training has stopped at 13^{th} epoch and achieved a training accuracy of 95% and validation accuracy of 98.8%. The figure below shows the transfer learning model using Xception:

Fig. 4. Xception with top layers added



2) **Fine tuning**:
   - For fine tuning the model, the entire base model is unfrozen by setting trainable equal to true.
   - The model is recompiled and then the training is performed with a very low learning rate (1e-5) for 20 epochs and achieved a training accuracy of 99.6% and a validation accuracy of 98.3%.

IV. Results and Discussion

The table below shows the fine tuned model evaluation performance on train,validation and test data sets:

	test	validation	train
loss	1.77	0.112	0.015
accuracy	0.826	0.983	0.996

The fine tuned Xception model has achieved an accuracy of 82.6% on test data, where as simple CNN is able to achieve only 66% accuracy on test set. The graphs below shows the train,validation accuracy and train,validation loss of both models.

Fig. 5. Simple CNN

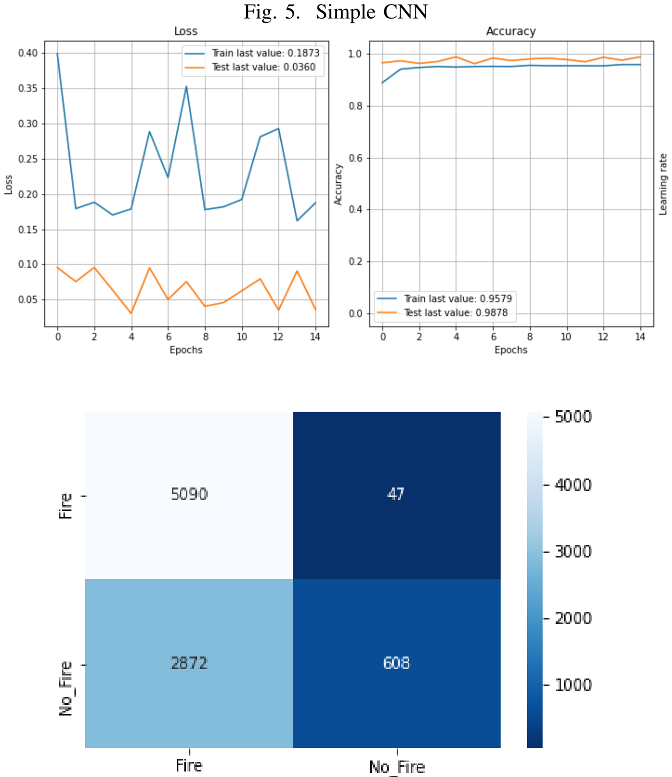


Fig. 6. Fine tuned Xception Model

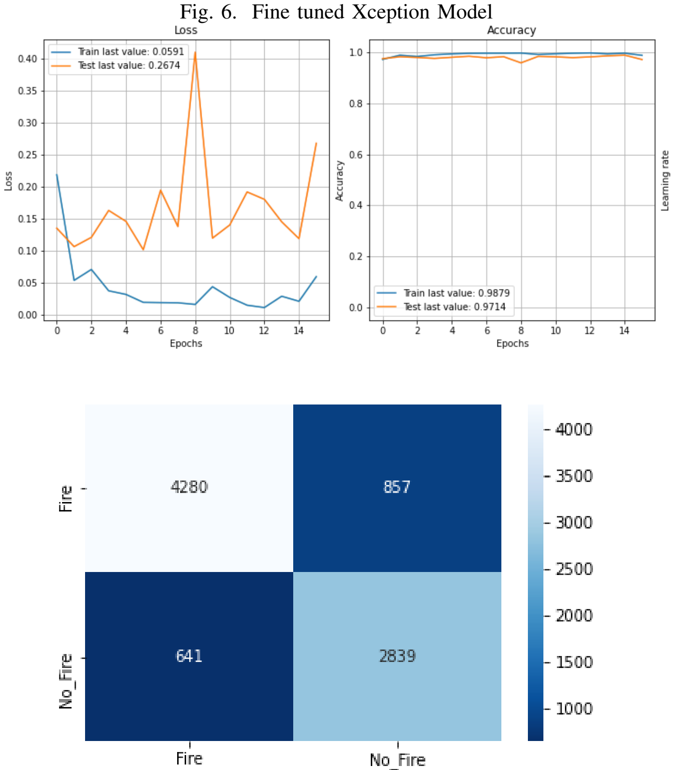


Fig. 7. Simple CNN Classification Report

Classification Report

	precision	recall	f1-score	support
fire	0.64	0.99	0.78	5137
no_fire	0.93	0.17	0.29	3480
accuracy			0.66	8617
macro avg	0.78	0.58	0.54	8617
weighted avg	0.76	0.66	0.58	8617

Fig. 8. Fine tuned Xception Model Classification Report

Classification Report

	precision	recall	f1-score	support
fire	0.87	0.83	0.85	5137
no_fire	0.77	0.82	0.79	3480
accuracy			0.83	8617
macro avg	0.82	0.82	0.82	8617
weighted avg	0.83	0.83	0.83	8617

Accuracy alone can't be used to judge model performance since the data is quite imbalanced. F1 score is good metric to evaluate performance of model as it combines both accuracy and precision. The equations for accuracy and F1 score are given below:

$$Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{1}$$

$$F_{measure} = \frac{2 * \text{precision} * \text{recall}}{\text{precision + recall}} \tag{2}$$

In case of simple CNN, the F1 score for fire class is 0.78 and 0.29 for no-fire class. Whereas the f1 score is found to be 0.85 and 0.79 for fire and no-fire class respectively using Xception model. Precision or True Positive Rate for fire class is increased drastically from 0.64 to 0.87 using fine tuned model.Also simple CNN predicted 2872 instances as no fire when there is a fire, where as fine tuned model has made 641 mistakes in detecting fire (False Negatives). This shows that the model is trained well to detect fire in most instances.

V. Conclusion

Availability of deep learning models and increase in computational power is motivating researchers to develop more advanced and robust systems to detect fire. This motivated us to leverage the power of transfer learning technique using deep learning models to create a fire detection model that can distinguish fire from non-fire images in real time.

Because of their deep architecture and the fact that they are trained on millions of images deep learning models such as Xception, are extremely useful even when the data available for training is less. It is evident from the work done in this paper that transfer learning is very effective compared to training the whole model from scratch or using traditional CNN approach alone.

Future work includes improving the model performance and also to locate and detect fire in the image.

REFERENCES

[1] https://www.nifc.gov/fireInfo/fireInfo$_{s}tats_{t}otalFires.html$

[2] Craig A. Matticks, MD, Jay J. Westwater, MD, Harvey N. Himel, MD, Raymond F. Morgan, MD, F. Edlich, MD, PhD, Health Risks to Fire Fighters, The Journal of Burn Care Rehabilitation, Volume 13, Issue 2, March-April 1992, Pages 223–235, https://doi.org/10.1097/00004630-199203000-00010

[3] https://www.forestresearch.gov.uk/research/climate-change-impacts/forest-fires-and-climate-change/

[4] Zivanovic, Stanimir Zigar, Darko Krstić, Dejan. (2013). The Role of Early Detection of Forest Fire in Environmental Protection, Safety Engineering. Safety Engineering. 3. 93-97.10.7562/SE2013.3.02.06,.

[5] D. Kinaneva, G. Hristov, J. Raychev and P. Zahariev, "Application of Artificial Intelligence in UAV platforms for Early Forest Fire Detection," 2019 27th National Conference with International Participation (TELECOM), Sofia, Bulgaria, 2019, pp. 50-53, doi: 10.1109/TELECOM48729.2019.8994888.

[6] M. Novák et al., "Multisensor UAV System for the Forest Monitoring," 2020 10th International Conference on Advanced Computer Information Technologies (ACIT), Deggendorf, Germany, 2020, pp. 293-296, doi: 10.1109/ACIT49673.2020.9208993.

[7] V. Sherstjuk, M. Zharikova and I. Sokol, "Forest Fire-Fighting Monitoring System Based on UAV Team and Remote Sensing," 2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO), Kyiv, UKraine, 2018, pp. 663-668, doi: 10.1109/ELNANO.2018.8477527.

[8] N. Kalatzis, M. Avgeris, D. Dechouniotis, K. Papadakis-Vlachopapadopoulos, I. Roussaki and S. Papavassiliou, "Edge Computing in IoT Ecosystems for UAV-Enabled Early Fire Detection," 2018 IEEE International Conference on Smart Computing (SMARTCOMP), Taormina, Italy, 2018, pp. 106-114, doi: 10.1109/SMARTCOMP.2018.00080.

[9] V. Sherstjuk and M. Zharikova, "Fire-Front Recognition in UAV-Based Forest-Fire Monitoring System Using Fuzzy Rough Soft Sets," 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON), Lviv, Ukraine, 2019, pp. 1091-1096, doi: 10.1109/UKRCON.2019.8879829.

[10] M. Erdelj, E. Natalizio, K. R. Chowdhury and I. F. Akyildiz, "Help from the Sky: Leveraging UAVs for Disaster Management," in IEEE Pervasive Computing, vol. 16, no. 1, pp. 24-32, Jan.-Mar. 2017, doi: 10.1109/MPRV.2017.11.

[11] . T. Celik, H. Ozkaramanli and H. Demirel, "Fire Pixel Classification using Fuzzy Logic and Statistical Color Model," 2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07, 2007, pp. I-1205-I-1208, doi: 10.1109/ICASSP.2007.366130.

[12] S. Frizzi, R. Kaabi, M. Bouchouicha, J. M. Ginoux, E. Moreau, and F. Fnaiech.Convolutional neural network for video fire and smoke detection. In IECON 2016 -42nd Annual Conference of the IEEE Industrial Electronics Society, pages 877–882,Oct 2016.

[13] C.Tao, J.Zhang, and P.Wang. Smoke detection based on deep convolutional neural networks. In 2016 International Conference on Industrial Informatics - Computing Technology, Intelligent Technology, Industrial Information Integration (ICIICII),pages 150–153, Dec 2016.

[14] Bc. Tomas Polednik. Detection of fire in images and video using cnn. Excel@FIT,2015.

[15] Alireza Shamsoshoara, Fatemeh Afghah, Abolfazl Razi, Liming Zheng, Peter Fulé, Erik Blasch, November 19, 2020, "The FLAME dataset: Aerial Imagery Pile burn detection using drones (UAVs)", IEEE Dataport, doi: https://dx.doi.org/10.21227/qad6-r683.

[16] C. Li and Y. Bai, "Fire Flame Image Detection Based on Transfer Learning," 2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS), 2018, pp. 370-374, doi: 10.1109/CCIS.2018.8691171.

[17] A. Bari, T. Saini and A. Kumar, "Fire Detection Using Deep Transfer Learning on Surveillance Videos," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021, pp. 1061-1067, doi: 10.1109/ICICV50876.2021.9388485.

[18] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convo lutions," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 1800-1807, doi: 10.1109/CVPR.2017.195.