Applying artificial intelligence models for the automatic forest fire detection

Ismail El madafri
Escola d'Enginyeria Barcelona Est. Departament d'Enginyeria Gràfica i de Disseny

Marta Peña
Escola Tècnica Superior d'Enginyeria Industrial de Barcelona. Departament de Matemàtiques

Noelia Olmedo-Torre
Escola d'Enginyeria Barcelona Est. Departament d'Enginyeria Gràfica i de Disseny

Abstract

Throughout the past decade, the development of Artificial Intelligence-based devices for the automatic detection of early stage forest fires has been a growing focus. Computer Vision techniques are well-suited for this problem due to the distinctive visual characteristics of forest fires. The effectiveness of several Artificial Intelligence algorithms in a binary classification problem involving fire/non-fire images was assessed by comparing them using a publicly available dataset. The benchmark dataset was used to both train and evaluate the models. An optimization method was employed to train the Artificial Intelligence algorithms, resulting in a higher performance than that previously achieved by studies on the same dataset.

Keywords: Forest fire detection, Sustainable management, Graphic-based intelligent surveillance, Deep Learning, Transfer Learning, Models optimization.

1. Introduction

Investment in forests offers a variety of solutions to global change, biodiversity decline, and various other environmental issues due to their role in the carbon cycle, natural disaster mitigation and human livelihoods [1-3]. The Intergovernmental Panel on Climate Change Sixth Assessment Report (AR6) confirms that climate change is causing increased stress on forest ecosystems and exacerbating the occurrence and intensity of damaging wildfires [4]. Recent technological advances have led to an increased demand for the automated detection of forest fires at an early stage in order to mitigate their destructive potential and promote a sustainable management of the forested areas. The early warning, remote fire parameter acquisition and monitoring are crucial to wildfire management.

Over the past decade, the development of cloud-based networks of scanning cameras has opened a promising avenue for the technological monitoring of forest fires. The increasing availability of forest data due to advancements in remote sensing platforms have provided a comprehensive range of information from all types of forested areas. For instance, Unmanned Aerial Vehicles (UAVs) have become a prominent source of forestry data due to their cost-effectiveness, light weight, user-friendliness, and maneuverability, allowing them to carry a variety of sensors that provide high spatial and temporal resolution. The combination of Artificial Intelligence and UAV-based forest fire data has recently seen increased research with the potential for high detection performance at a relatively low cost [5;6].

Deep Learning is a type of Artificial Intelligence that utilizes algorithms composed of interconnected layers of neurons, known as Deep Neural Networks, to extract features from input data. These neurons are trained via the backpropagation technique, which adjusts the parameters of the model, allowing for the development of increasingly complex predictive models [7]. The integration of Convolutional Neural Networks (CNNs), a type of Deep Learning algorithms, has enabled significant progress in Computer Vision classification tasks in the last decade. However, the acquisition of labeled images of forest fires is difficult due to technical and safety constraints, making the majority of DL studies that process forest fire images rely on private datasets sourced from the web, which are usually of limited size [8]. The lack of public benchmarks, which difficulties reproducing the methods, hinders the widespread implementation of these promising models. To ensure accurate progress, the use of common benchmark datasets is necessary to enable the objective comparison of the different approaches [9].

For the above mentioned considerations, the present work conducted experiments using the DeepFire dataset [10], a publicly available dataset for fire/non-fire 3-channel Red-Green-Blue (RGB) image classification, composed of a total of 1900 images of different sizes, with 950 each for fire and non-fire classes. RGB images are attractive for their low cost, ability to capture high resolution, and ease of processing, which are beneficial for a real-world deployment.

2. Objetive of the study

This study conducted a comparative analysis of the performance of four state-of-the-art CNN architectures (VGG19 [11], DenseNet121

[12], Inception-ResNet V2 [13] and MobilNet V2 [14]) on a binary-classification task involving fire and non-fire images from the DeepFire dataset. The Transfer Learning technique was employed to train the models, using pre-trained architectures on the ImageNet dataset [15]. This was followed by fine-tuning on the DeepFire dataset, allowing for rapid learning of features from the new dataset without the need for model initialization from scratch. The study applies various Transfer Learning approaches to a comparative analysis in order to analyze the effects of varying the portion of the base network (also called backbone) that is re-trained on the classification performance of models in a forest fire detection task. Previous research on the effects of Transfer Learning on classification performance of Deep Learning models applied on forest fire images has been limited and there has been no such experiment regarding the effect of the number of frozen layers on the whole performance in a forest fire detection task.

3. Materials and methods

3.1. The Dataset

This work used the DeepFire dataset [15], which consists of 1900 RGB images of various forested areas. The fire and non-fire classes each contain 950 images. The training data was split into a training set of 1216 images and a validation set of 304 images, representing 20 % of the total training set, to update the network weights and tune the hyper-parameters, respectively. The same repartition was used for all the considered networks to decrease the randomness effect. The proposed original test dataset was used to evaluate the models' performance on unseen data after the training process (see Table 1.).

Dataset	Training Dataset	Held-out validation Dataset	Test Dataset	Total
Fire	608	152	190	950
Non-fire	608	152	190	950
Total	1012	304	380	1900

Table 1. Fire and non-fire classes for each sub-dataset before data augmentation.

3.2. Transfer Learning scenarios

The effect of varying the number of frozen layers in the backbones models was assessed by exploring five different scenarios, ranging from completely freezing to completely training the backbone network. For each scenario, the non-frozen layers were jointly optimized with newly-stacked fully connected layers on top of them.

Fine-tuning scenario	Portion of frozen layers
Scenario 1	The model base is trained from scratch
Scenario 2	freezing 20 % of the model base
Scenario 3	freezing 50 % of the model base
Scenario 4	freezing 80 % of the model base
Scenario 5	freezing 100 % of the model base

Table 2. Portion of frozen layers for each scenario.

Table 2 provides the proportion of frozen layers for each Transfer Learning scenario.

3.3 Training the models

In this study, the Keras framework with TensorFlow backend was used to implement the Convolutional Neural Networks with GPU support for compiling source code. Images used were originally colored, of varied sizes and had each pixel normalized by dividing by 255. The images were then resized to the default size of each model (224×224 pixels for VGG19, MobileNet V2, and DenseNets architectures, and 299×299 pixels for InceptionResNet V2), with 32-sized mini-batches used to train all networks. The binary cross-entropy (also called the log loss) was employed as the loss function to measure the difference between the network predictions and the labels. Three stochastic optimization methods were tested on each model: Stochastic Gradient Descent (SGD), Root Mean Square Propagation (RMSProp), and Adaptive Moment Estimation (Adam). The training phase of the models was limited to a maximum of 50 epochs, with an early stopping configured with a patience of 5 and a minimum change in the loss of 1e-3. The optimal Learning Rate for each model was determined by testing four different Learning Rates (10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}). The combination of the Learning Rate and Optimizer with the highest accuracy on the held-out validation dataset was selected to evaluate it on the test dataset. Additionally, data augmentation was applied to the training dataset through random transformations (random horizontal flipping and random rotation) to make the training set more representative and reduce bias in the images. The BatchNormalization layers within the unfrozen portion of the imported backbone model were kept in inference mode to avoid any interference of the updated non-trainable parameters/ weights with the feature maps the model had learned. To reduce the dimensionality of the output matrix from the convolutional layers, a global average pooling layer was integrated, which computes the average value of each feature map. The output matrix was subsequently flattened into a vector, which was used as an input for the fully connected prediction layer, allowing the model to generate predictions based on the extracted features from the input data. Regularization techniques such as Dropout with a dropout rate of 0.2 were employed, meaning that 20 % of the neurons in each layer were randomly set to zero during each training iteration. To evaluate the overall performance of the models, accuracy, precision, recall and F1-score were computed.

4. Results

The performance of four Convolutional Neural Network architectures (VGG19, DenseNet121, InceptionResNet V2 and MobileNet V2) was evaluated in a comparative study using the DeepFire Benchmark dataset. Results demonstrated that InceptionResNet V2 achieved the highest average score. Further analysis revealed that a half re-trained InceptionResNet V2 base model, trained with Adam and a Learning Rate of 10^{-5}, achieved the highest score (100 %) across all four performance metrics considered. These findings are promising, given the relative representativeness of the DeepFire Benchmark dataset.

Beyond trying to find a highly performing model for the task at hand, this study aimed to assess the hypothesis that varying the portion of the trainable model base in a pre-trained Convolutional Neural Network may have a positive effect on the model's performance. The results of the experimental scenarios 2, 3 and 4 demonstrated an increased classification performance when compared to scenarios 1 and 5. The scores achieved on the test dataset support the conclusions drawn. The employed optimization method resulted in a higher performance than that previously achieved by studies on the same dataset.

5. Discussion

Intelligent surveillance systems utilizing Deep Learning technology have been proposed as a potential solution for the automated monitoring of forests in order to prevent forest fires. To assess the efficacy of such models, a comparative study was conducted to evaluate various CNN architectures on the DeepFire Benchmark dataset. The study findings showed that the InceptionResNet V2 network achieved the highest average score and the best performing model was found to be a half re-trained InceptionResNet V2 base model (Scenario 3) trained with Adam as an optimizer and 10^{-5} as a Learning Rate. This model achieved a 100 % score in all the four considered performance metrics. The findings of this study provide insights into the use of CNNs for the detection task of forest fire and suggest that such neural networks can be effectively integrated into an inference system for the purpose of forest sustainable monitoring.

Our experiments also evaluated the performance of various frozen layer configurations in Transfer Learning for enhancing the accuracy of classification. The results demonstrated that the employed optimization technique yielded higher performance than that reported in previous studies on the same dataset. Indeed, building upon the results of references [10] and [16], the present study seeks to demonstrate how the optimization of the number of frozen layers can improve the detection performance of the two models that had been proposed.

This work demonstrates the potential of using RGB images and Deep Learning for the practicalization of early forest fire detection systems. The results obtained indicate that the proposed approach could be a viable option for improving forest fire detection with the potential for providing an effective and cost-efficient solution.

References

[1] Rizvi, A.R., Baig, S., Barrow, E., Kumar, C. (2015). *Synergies between Climate Mitigation and Adaptation in Forest Landscape Restoration*. Gland, Switzerland: IUCN.

[2] https://www.worldbank.org/en/topic/forests *Investing in forests* (Visited the 27/12/2022)

[3] https://sdgs.un.org/topics/forests United Nations, The Agenda for Sustainable Development, https://www.un.org/ (Visited the 27/12/2022)

[4] Seneviratne, S.I. et al. Weather and Climate Extreme Events in a Changing Climate. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, 1513–1766. https://doi.org/10.1017/9781009157896.013

[5] Bouguettaya, A. et al. (2022). A review on early wildfire detection from unmanned aerial vehicles using deep learning-based computer vision algorithms. *Sign. Proc.*, 190, 108309. https://doi.org/10.1016/j.sigpro.2021.108309

[6] Kattenborn, T., Leitloff, J., Schiefer, F., Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in vegetation remote sensing, ISPRS. *Journal of Photogrammetry and Remote Sensing*, 173, 24-49. ISSN 0924-2716. https://doi.org/10.1016/j.isprsjprs.2020.12.010

[7] LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. *Nature* 521, 436-444. https://doi.org/10.1038/nature14539

[8] Diez, Y., Kentsch, S., Fukuda, M., Caceres, M. L. L., Moritake, K., & Cabezas, M. (2021). Deep Learning in Forestry Using UAV-Acquired RGB Data: A Practical Review. *Remote Sensing*, 13(14), 2837. https://doi.org/10.3390/rs13142837

[9] Hutson, M. (2018). Artificial intelligence faces reproducibility crisis. Science, 359(6377), 725-726. https://doi.org/10.1126/science.359.6377.725

[10] Khan, A., Hassan, B., Khan, S., Ahmed, R., & Abuassba, A. (2022). DeepFire: A Novel Dataset and Deep Transfer Learning Benchmark for Forest Fire Detection. *Mobile Information Systems*, Article ID 5358359, 14. https://doi.org/10.1155/2022/5358359

[11] Simonyan, K. et al. Very deep convolutional networks for large-scale image recognition. *arXiv* preprint https://doi.org/10.48550/arXiv.1409.1556

[12] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L-C. (2018). Mobilenetv2: Inverted Residuals and Linear Bottlenecks. *arXiv* preprint arXiv:1801.04381. https://doi.org/10.48550/arXiv.1801.04381

[13] Huang G. et al. Densely Connected Convolutional Networks. *arXiv*:1608.0693 https://doi.org/10.48550/arXiv.1608.06993

[14] Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2016). Inception-v4, inception-resnet and the impact of residual connections on learning. *arXiv* preprint https://doi.org/10.48550/arXiv.1602.07261

[15] Deng, J., Dong, W., Socher, R., Li, L. -J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, 248-255. https://doi.org/10.1109/CVPR.2009.5206848

[16] Khan, S., & Khan, A. (2022). FFireNet: Deep Learning Based Forest Fire Classification and Detection in Smart Cities. *Symmetry*, 14(10), 2155. https://doi.org/10.3390/sym14102155