



HARNESSING DEEP LEARNING FOR WILDFIRE RISKS PREDICTION: A NOVEL APPROACH

Hoang Anh Duc

Department of Software Engineering, Faculty of
Information Technology, Hanoi University of Mining and
Geology
Hanoi, Vietnam

Le Van Hung

Department of Software Engineering, Faculty of
Information Technology, Hanoi University of Mining and
Geology
Hanoi, Vietnam

Nguyen Thi Huu Phuong

Department of Software Engineering, Faculty of Information Technology,
Hanoi University of Mining and Geology
Hanoi, Vietnam

Abstract: This article presents a pioneering approach for predicting wildfires risks using deep learning techniques. By combining convolutional neural networks (CNNs), recurrent neural networks (RNNs) and Adaptive Moment Estimation (ADAM), our framework analyses geospatial and environmental data to capture the intricate dynamics of disasters. Our model integrates satellite imagery, climate data, socioeconomic factors, and historical records to accurately assess risks. Leveraging transfer learning, we optimize training efficiency with pre-trained models. Extensive experiments demonstrate the superior performance of our deep learning framework compared to traditional methods. With its ability to enable proactive planning and decision-making, our approach strengthens disaster preparedness and response strategies. This research represents a significant advancement in utilizing deep learning for predicting wildfires risks, paving the way for further innovations in this vital field.

Keywords: Deep learning; CNNs; RNNs; ADAM; Wildfires; Geospatial data analysis.

I. INTRODUCTION

The escalating prevalence and devastating consequences of wildfires on a global scale have emphasized the urgent need for innovative, accurate, and timely methods of risk prediction. In this article, we address this pressing issue by introducing a cutting-edge approach to wildfire risk prediction using advanced deep learning techniques. Our methodology harnesses the potential of convolutional neural networks (CNNs), recurrent neural networks (RNNs), and Adaptive Moment Estimation (ADAM) within an integrated model, enabling us to unravel the intricate dynamics of these catastrophic events. We delve into the details of how our deep learning model effectively utilizes a diverse range of data sources, including satellite imagery, climate data, socioeconomic factors, and historical records, to precisely assess wildfire risks. Importantly, our strategy optimizes training efficiency through transfer learning, leveraging the power of pre-trained models to enhance predictive performance. Through extensive experimental evaluations, we provide compelling evidence of the superior capabilities of our model compared to traditional methods. We emphasize the significant role that deep learning can play in bolstering disaster preparedness and response strategies. The insights derived from this research represent a notable advancement in the application of deep learning techniques to wildfire risk prediction, and they pave the way for future exploration and advancements in this critical field.

II. METHODOLOGY

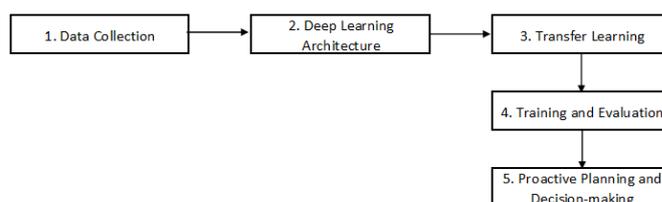


Figure 1 Proposed methodology

In this study, we explore the application of deep learning methods, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and the ADAM optimization algorithm, for predicting wildfires. The processes involved in the study are depicted in Figure 1. The steps involved in determining them are as follows:

Step 1: Data Collection: Gather a diverse range of data types, including satellite imagery, climate data, socioeconomic factors, and historical records, to ensure a comprehensive understanding of the factors influencing wildfire risks.

Step 2: Deep Learning Architecture: Develop an integrated model that combines convolutional neural networks (CNNs), recurrent neural networks (RNNs), and Adaptive Moment Estimation (ADAM). This hybrid architecture enables effective capture and analysis of the complex dynamics associated with wildfires.

Step 3: Transfer Learning: Implement transfer learning by utilizing pre-trained models to initialize the deep learning framework. This approach leverages learned features and optimizes training efficiency.

Step 4: Training and Evaluation: Train the deep learning model using the integrated architecture and the collected

dataset. Conduct extensive experiments on real-world datasets to evaluate the model's performance. Compare the results obtained by the deep learning framework with those of traditional methods to highlight its superiority in predicting wildfire risks.

Step 5: Proactive Planning and Decision-making: Emphasize the importance of utilizing the trained model for proactive planning and decision-making in disaster preparedness and response strategies. The timely and accurate predictions provided by the model contribute to strengthening these critical aspects of wildfire management.

Overall, the methodology involves leveraging deep learning techniques, diverse data sources, transfer learning, and rigorous experimentation to develop an advanced model for wildfire risk prediction.

III. THE STUDY AREA AND DATA USED

A. Description of the study area

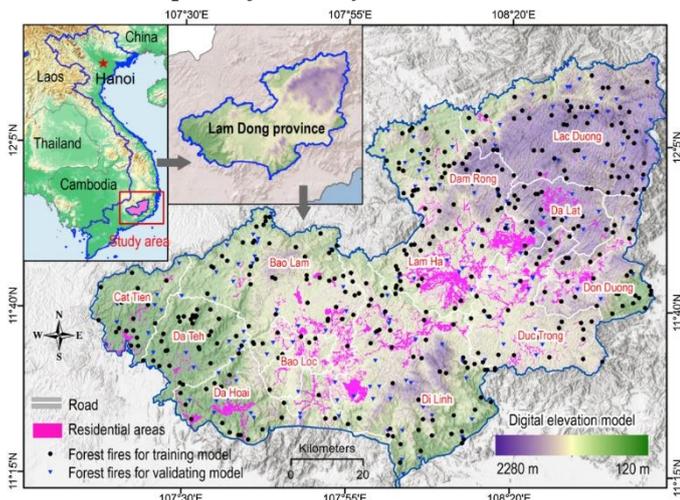


Figure 2 Location of the study area and historical wildfires, source [1]

Lam Dong province, as illustrated in Figure 2, is situated in the southern part of Vietnam's Central Highland region. It is positioned between latitudes 11° 12' 00" N and 12° 15' 00" N, and longitudes 107° 15' 00" E and 108° 45' 00" E. Covering an area of 9805.4 km², the province boasts a diverse topography. The elevation ranges from 120 m to 2280 m above sea level, with an average of 907.6 m and a standard deviation of 392.1 m.

The climate in Lam Dong province is influenced by a tropical monsoon, characterized by moderate temperatures and high humidity. The local climate varies based on altitude and can be classified into two distinct seasons: the rainy season, which spans from May to November, and the dry season, lasting from December to April. Temperature variations are significant across the region, with temperatures ranging from 18°C to 25°C, resulting in mild and cool weather conditions. During the rainy season, rainfall contributes to about 90% of the total annual precipitation, with values ranging from 1600 to 2700 mm per year. The average relative humidity throughout the year ranges between 85% and 87%.

Forest coverage accounts for approximately 60% of the total study area in Lam Dong province [2], while agricultural land and populated areas make up approximately 28% and 6% respectively. The remaining land covers various other types. The dominant tree species include *Suzygium*, *Dipterocarpus*, *Anisopkeracochinchinensia*, and *Schima superba* Gardner & Champ (found between 1000 and 1300 m), *Pinus merkusii*

(between 600 and 1000 m), *Pinus khasya* (>1000 m), and *Dipterocarpus obtusifolius* and *Shorea obtusa* (at 1300 m). According to Dien and colleagues [2], deforestation and forest degradation have been significant threats to the forest cover in the province during the period from 1990 to 2010. Despite a considerable increase of 81.7% in plantation forests (representing 5.1% of the total study area), the overall forest area has experienced a reduction of 118067.9 ha (equivalent to 12.04% of the total study area). Broadleaf forest, bamboo forest, and coniferous forest have been particularly affected, with reductions of 30.5%, 37.1%, and 28.2% respectively. These degradation processes encompass 17% of the total study area, with causes of forest loss including fire, illegal logging, land conversion, and inadequate management [2]. The increasing population growth has exerted significant pressure on forest resources, driven by the rising demand for residential and production lands.

B. Historical wildfires

Due to the reliance on historical forest fires and their associated ignition factors for developing prediction models of forest fire susceptibility [3][4], it is essential to create a forest fire inventory map. In this particular study, a comprehensive forest fire inventory map was compiled, encompassing a total of 540 historical fire locations. These fire incidents, which took place in 2013, were provided by the Department of Forest Protection under the Ministry of Agriculture and Rural Development of Vietnam (2016) and can now be accessed at <http://www.kiemlam.org.vn/firewatchvn/>. This national database serves as the official source of information on forest fires in Vietnam.

For the purpose of this analysis, only the forest fires that occurred in 2013 were chosen. This specific year was marked by the most severe drought in the study area over the past three decades[5]. Notably, our examination of these fire locations revealed a significant occurrence of forest fires in March, accounting for 39.1% of the total. In contrast, no forest fires were reported for the months of July, August, September, October, and December. Further detailed statistical analysis regarding the temporal distribution of forest fires is presented in Table 1.

Table 1 Analysis of the temporal occurrence of forest fires in this study

No	Forest fires (%)	Month
1	12.8	Jan
2	19.8	Feb
3	39.1	Mar
4	20.7	Apr
5	5	May
6	1.5	Jun
7	1.1	Nov
8	0	Jul, Aug, Sep, Oct, Dec

C. Factors influencing forest fire ignition

Choosing the suitable ignition factors for forest fire modeling is a crucial matter that impacts the accuracy of the resulting prediction models. In the study conducted by Bui *et al.* [1], it has been demonstrated that factors such as slope, elevation, aspect, land use, NDVI (Normalized Difference Vegetation Index), distance to road, distance to residence area, temperature, wind speed, and rainfall play a significant role in determining wildfire susceptibility.

In this study, the DEM (Digital Elevation Model) of Lam Dong province was utilized to extract key factors such as slope, aspect, and elevation. Additionally, elevation was chosen due to its potential influence on precipitation, temperature,

humidity, and evapotranspiration [6], was also included in the analysis. The aspect factor, crucial for forest fire modeling [7], is significant because it influences soil moisture and wind speed—two elements that significantly affect fire behavior [8].

To enhance the understanding of forest fire susceptibility, it is essential to consider human land use activities as potential ignition sources [9][10]. Additionally, the Normalized Difference Vegetation Index (NDVI) is a crucial factor in forest fire modeling, as it reflects the health status of vegetation [11], which serves as a proxy for fuel load distribution [12]. In this study, NDVI was computed using Landsat-8 Operational Land Imagery with a resolution of 30 m, obtained from the USGS archive (available at <http://earthexplorer.usgs.gov>), employing the following equation:

$$NDVI = (NIR-RED)/(NIR+RED)$$

The Normalized Difference Vegetation Index (NDVI) was generated using the near-infrared band (NIR) (0.76–0.90 μm , Band 4) and the red band (RED) (0.63–0.69 μm , Band 3).

In Vietnam, many forest fires are linked to anthropogenic activities such as grass burning, hunting with fire, and forest exploitation [13]. To account for this, the analysis included the distance to roads and residential areas. The road network was extracted from national topographic maps at a scale of 1:50,000, and the distance to road map was created by calculating the Euclidean distance from the road lines using the buffer tool in ArcGIS 10.2. Residential areas in the Lam Dong province were extracted from the land use map and used to construct the residential area map using the same buffer tool.

Climatic factors, including air temperature, wind speed, and rainfall, have been found to influence the severity and frequency of forest fires [14][15] as well as soil moisture and drought [16]. Thus, temperature, wind, and rainfall data for the Lam Dong province in 2013 were obtained from the Climate Forecast System Reanalysis available at <https://www.ncdc.noaa.gov/>.

IV. CNNs, RNNs ALGORITHMS AND ADAM OPTIMIZER

A. CNNs Algorithm

Convolutional Neural Networks (CNNs) [17] are a powerful class of deep learning algorithms extensively utilized for a wide range of image and video processing tasks, encompassing object recognition, image classification, and object detection. CNNs are specifically tailored to effectively analyze and extract meaningful information from data with grid-like structures, such as images. They achieve this by leveraging the concept of convolution, which enables the network to detect local patterns and features within the input data. This characteristic makes CNNs highly effective in automatically learning and identifying complex visual patterns, making them a valuable tool in computer vision applications.

The fundamental principle behind CNNs is to apply convolutional layers, pooling layers, and fully connected layers to extract meaningful features from the input data. The convolutional layers employ learnable filters or kernels to scan the input data, performing convolutions that aid in detecting local patterns or features. These convolutions yield feature maps that capture distinct aspects of the input data.

Here is an illustrative example of the CNN algorithm in pseudo code:

Input: Training data (images) and corresponding labels

Step 1. Initialize the CNN architecture:

- Specify the number and size of convolutional filters.
- Determine the number and size of pooling layers.
- Define the number and size of fully connected layers.

Step 2. Forward propagation:

- Convolutional layers:
- Apply convolutional filters to the input images.
- Introduce non-linearity by applying an activation function (e.g., ReLU).

- Pooling layers:

- Reduce the spatial dimensions of the feature maps.
- Retain important features through down sampling (e.g., max pooling).

- Flatten the feature maps into a 1D vector.

- Fully connected layers:

- Connect each neuron to every neuron in the previous layer.
- Introduce non-linearity by applying an activation function.

Step 3. Backpropagation:

- Calculate the loss between predicted and actual labels.

- Compute the gradients of the loss with respect to the network parameters.

- Update the parameters using gradient descent or other optimization algorithms.

Step 4. Repeat steps 2 and 3 for a defined number of epochs or until convergence.

Step 5. Evaluate the trained CNN:

- Apply the trained CNN to unseen data.

- Assess performance using appropriate evaluation metrics (e.g., accuracy).

In this research, Convolutional Neural Networks (CNNs) were utilized through the Conv1D layer of Keras. Specifically, two Conv1D layers were added to the model, each with 64 filters and a kernel size of 3. Additionally, the model employed the Rectified Linear Unit (ReLU) activation function in both of these Conv1D layers.

B. RNNs Algorithm

Recurrent Neural Networks (RNNs) [18] is a class of artificial neural networks designed to process sequential data by incorporating feedback connections. Unlike feedforward neural networks, RNNs have connections that allow information to be passed from previous time steps to the current time step, enabling them to capture temporal dependencies and handle input sequences of varying lengths.

Here is an example of the pseudo code for the basic operation of an RNN:

Step 1: Initialize weights, biases, and the hidden state.

Step 2: For each input sequence and time step:

- Combine input with the previous hidden state.
- Compute the activation of the recurrent unit.
- Update the hidden state.

Step 3: Compute the output using the final hidden state.

Step 4: Calculate the loss based on the predicted output and target.

Step 5: Backpropagate the error through time to update weights and biases.

In this research, the authors employed Recurrent Neural Networks (RNNs) using the LSTM layer of Keras. LSTM (Long Short-Term Memory) is an improved variant of RNN designed to address the gradient vanishing/exploding problem encountered during RNN training.

C. Adam optimizer Algorithm

The Adam optimizer algorithm [19] is a widely used optimization method in deep learning. It combines the advantages of both Adaptive Gradient Algorithm (AdaGrad) [20] and Root Mean Square Propagation (RMSprop) [21] algorithms. Adam stands for Adaptive Moment Estimation, and it is known for its efficiency in adjusting learning rates for different parameters during the training process. This adaptive learning rate optimization algorithm helps accelerate convergence and improve the overall performance of neural networks.

Here is the pseudocode for the Adam optimization algorithm. This algorithm uses estimates of first and second

moments of gradient to adapt the learning rate for different weights.

Initialize parameters:

- learning rate: alpha (usually set to 0.001)
- first moment vector: m (initialize as zero vector)
- second moment vector: v (initialize as zero vector)
- weight parameters: theta
- decay rates for moment estimates: beta1, beta2 (usually set to 0.9 and 0.999 respectively)
- small constant for numerical stability: epsilon(usually set to 1e-8)
- time step: t (initialize as 0)

For each iteration:

t = t + 1

Get gradients: g = ComputeGradients(theta)

Update biased first moment estimate: m = beta1*m+(1-beta1)*g

Update biased second raw moment estimate:

v=beta2*v+(1-beta2)*g*g

Correct bias in first moment: m_hat= m / (1 - beta1^t)

Correct bias in second moment: v_hat = v / (1 - beta2^t)

Update parameters: theta=theta-alpha*m_hat/(sqrt(v_hat)+epsilon)

End For

V. EXPERIMENT RESULT

A. Dataset

In order to evaluate the effectiveness of the proposed method, the author chose a dataset consisting of 755 training data points and 324 validation data points stored in two *.arff files. The data included parameters: slope, aspect, elevation, land use, NDVI, distance to road, distance to residence area, temperature, wind speed, rainfall and forest fires resulting in sample point.

B. Results

The performance of the proposed method was then compared with that of using only CNNs and RNNs algorithms with the Adam optimizer. Therresults are presented below:

Validation accuracy: 78.40%
Model: "sequential"

Layer (type)	Output Shape	Param #
conv1d (Conv1D)	(None, 8, 64)	256
dropout (Dropout)	(None, 8, 64)	0
conv1d_1 (Conv1D)	(None, 6, 64)	12352
dropout_1 (Dropout)	(None, 6, 64)	0
flatten (Flatten)	(None, 384)	0
dense (Dense)	(None, 1)	385

Total params: 12,993
Trainable params: 12,993
Non-trainable params: 0

Figure 3Result of the method using CNNs with Adam optimizer.

Validation accuracy: 70.99%
Model: "sequential"

Layer (type)	Output Shape	Param #
flatten (Flatten)	(None, 10)	0
dense (Dense)	(None, 64)	704
dropout (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 64)	4160
dropout_1 (Dropout)	(None, 64)	0
dense_2 (Dense)	(None, 1)	65

Total params: 4,929
Trainable params: 4,929
Non-trainable params: 0

Figure 4Result of the method using RNNs with Adam optimizer.

Validation accuracy: 87.04%
Model: "sequential"

Layer (type)	Output Shape	Param #
conv1d (Conv1D)	(None, 8, 64)	256
conv1d_1 (Conv1D)	(None, 6, 64)	12352
conv1d_2 (Conv1D)	(None, 4, 128)	24704
lstm (LSTM)	(None, 100)	91600
dropout (Dropout)	(None, 100)	0
dense (Dense)	(None, 1)	101

Total params: 129,013
Trainable params: 129,013
Non-trainable params: 0

Figure 5Result after running model using combination of CNNs, RNNs and Adam

C. Discussion

From the comparison provided, we can see that the best performance was achieved by combining CNNs (Convolutional Neural Networks), RNNs (Recurrent Neural Networks, specifically LSTM in this case), and using the Adam optimizer. This model gave a validation accuracy of 87.04%.

On the other hand, when only CNNs or only RNNs were used along with Adam optimizer, the accuracy was lower. For the model with only CNNs, the validation accuracy was 78.4%, and for the one with only RNNs, the accuracy was 70.99%.

Combining both CNNs and RNNs in a model could potentially allow the model to capture both spatial and temporal patterns in the data, leading to better performance.

However, also consider the trade-off in computational cost and model complexity. The combined CNN-RNN model has significantly more parameters (129,013) compared to the models using only CNNs (12,993) or RNNs (4,929), which would require more computational resources to train and could potentially overfit if not regularized properly.

These results show that model selection should not only be based on performance metrics but also other factors like computational resources, model interpretability, and overfitting.

VI. CONCLUSIONS

In drawing this piece to a close, the article has successfully presented a groundbreaking and innovative approach to predicting the risk of wildfires utilizing sophisticated deep learning techniques. This innovative framework leverages the robust capabilities of Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and the Adaptive Moment Estimation (ADAM) optimizer to analyze a comprehensive range of data effectively. This spectrum of data, inclusive of satellite imagery, climate data, socioeconomic factors, and historical records, contributes to generating highly accurate risk assessments for wildfires.

The performance of our model, as demonstrated through extensive and meticulous experiments, exhibits a noticeable superiority when compared to traditional methods that have previously been employed in this field. Furthermore, the model's architecture and training efficiency have been significantly enhanced through the strategic application of transfer learning, utilizing pre-trained models to reduce training time and increase prediction accuracy.

One of the pivotal strengths of our model lies in its capacity to enable more proactive planning and informed decision-making. This substantially enhances disaster preparedness and response strategies, making it a potent tool for disaster management personnel and policy makers.

This research piece serves as a substantial contribution to the field of deep learning applications in disaster risk

prediction, especially pertaining to wildfires. However, it is not only the result but also the journey that matters. The methodologies adopted and findings presented within this study have the potential to serve as a significant cornerstone for future research initiatives. They pave the way for exploration into leveraging similar advanced AI techniques for risk prediction across a multitude of fields. This work, therefore, is a significant stride forward in our journey towards a safer and more resilient future.

VII. ACKNOWLEDGMENT

This research is sponsored by Hanoi University of Mining and Geology under the research project number T22-04, led by MSc. Hoang Anh Duc: "A study on building extensions in the ArcGIS Pro Environment for disaster risk assessment based on Deep Learning methods."

VIII. REFERENCES

- [1] D. T. Bui, Q.-T. Bui, Q.-P. Nguyen, B. Pradhan, H. Nampak, and P. T. Trinh, "A hybrid artificial intelligence approach using GIS-based neural-fuzzy inference system and particle swarm optimization for forest fire susceptibility modeling at a tropical area," *Agricultural and Forest Meteorology*, vol. 233, pp. 32-44, 2017. [Online]. Available: <https://doi.org/10.1016/j.agrformet.2016.11.002> Accessed: Dec. 22, 2022.
- [2] V. T. Dien, P. N. Bay, P. Stephen, T. V. Chau, A. Grais, and S. Petrova, "Land Use, Forest Cover Change and Historical GHG Emission from 1990 to 2010, Lam Dong province, Vietnam," USAID, LEAF Hanoi, 2013.
- [3] P. E. Higuera, J. T. Abatzoglou, J. S. Littell, and P. Morgan, "The Changing Strength and Nature of Fire-Climate Relationships in the Northern Rocky Mountains, U.S.A., 1902-2008," *Published June 26, 2015*. [Online]. Available: <https://doi.org/10.1371/journal.pone.0127563>. Accessed: Dec. 22, 2022.
- [4] D. Tien Bui, K.-T. Le, V. Nguyen, H. Le, and I. Revhaug, "Tropical Forest Fire Susceptibility Mapping at the Cat Ba National Park Area, Hai Phong City, Vietnam, Using GIS-Based Kernel Logistic Regression," *Remote Sensing*, vol. 8, no. 4, p. 347, Apr. 2016, doi: 10.3390/rs8040347.
- [5] N. L. Dan, N. T. Hieu, and V. T. T. Lan, "Drought, desertification in Tay Nguyen territory associated with climate change scenarios," *J. Earth Sci.*, vol. 35, no. 4, pp. 310-317, 2014.
- [6] J. C. Verde and J. L. Zêzere, "Assessment and validation of wildfire susceptibility and hazard in Portugal," *Nat. Hazards Earth Syst. Sci.*, vol. 10, no. 3, pp. 485-497, 2010. [Online]. Available: <https://nhess.copernicus.org/articles/10/485/2010/>. [Accessed: 01/01/2023].
- [7] A. Camp, C. Oliver, P. Hessburg, and R. Everett, "Predicting late-successional fire refugia pre-dating European settlement in the Wenatchee Mountains," *Forest Ecology and Management*, vol. 95, no. 1, pp. 63-77, 1997. [Online]. Available: [https://doi.org/10.1016/S0378-1127\(97\)00006-6](https://doi.org/10.1016/S0378-1127(97)00006-6). Accessed on: Jan. 1, 2023.
- [8] D.A. Schmidt, A.H. Taylor, and C.N. Skinner, "The influence of fuels treatment and landscape arrangement on simulated fire behavior, Southern Cascade range, California," *Forest Ecology and Management*, vol. 255, no. 8, pp. 3170-3184, 2008.
- [9] M. Huesca, J. Litago, A. Palacios-Orueta, F. Montes, A. Sebastián-López, and P. Escribano, "Assessment of forest fire seasonality using MODIS fire potential: A time series approach," *Agricultural and Forest Meteorology*, vol. 149, no. 11, pp. 1946-1955, 2009. [Online]. Available: <https://doi.org/10.1016/j.agrformet.2009.06.022>. Accessed: Jan. 1, 2023.
- [10] D. Nepstad, C. Stickler, B. Filho, and F. Merry, "Interactions among Amazon Land Use, Forests and Climate: Prospects for a Near-Term Forest Tipping Point," *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, vol. 363, pp. 1737-1746, 2008. [Online]. Available: <https://doi.org/10.1098/rstb.2007.0036>. Accessed: Dec. 22, 2022.
- [11] S. Bajocco, E. Dragoz, I. Gitas, D. Smiraglia, L. Salvati, "Mapping Forest Fuels through Vegetation Phenology: The Role of Coarse-Resolution Satellite Time-Series," *PLOS ONE*, vol. 10, no. 3, pp. e0119811, 2015. [Online]. Available: <https://doi.org/10.1371/journal.pone.0119811>. Accessed: Dec. 22, 2022.
- [12] K. Yi, H. Tani, J. Zhang, M. Guo, X. Wang, and G. Zhong, "Long-Term Satellite Detection of Post-Fire Vegetation Trends in Boreal Forests of China," *Remote Sensing*, vol. 5, no. 12, pp. 6938-6957, 2013. [Online]. Available: <https://doi.org/10.3390/rs5126938>. Accessed: Dec. 22, 2022.
- [13] H. Le, T. Nguyen, K. Lasko, S. Ilavajhala, K. Vadrevu, and C. Justice, "Vegetation fires and air pollution in Vietnam," *Environmental Pollution*, vol. 195, pp. 10-19, 2014. [Online]. Available: <https://doi.org/10.1016/j.envpol.2014.07.023>. Accessed: Dec. 22, 2022.
- [14] N. Gillett, A. Weaver, F. Zwiers, and M. Flannigan, "Detecting the effect of climate change on Canadian forest fires," *Geophysical Research Letters*, vol. 31, Sep. 2004, doi: 10.1029/2004GL020876. Accessed: Dec. 22, 2022.
- [15] M. Heimann and M. Reichstein, "Terrestrial ecosystem carbon dynamics and climate feedbacks," *Nature*, vol. 451, pp. 289-292, Jan. 2008. [Online]. Available: <https://www.nature.com/articles/nature06591>. Accessed: Dec. 22, 2022.
- [16] B. Zaitchik, J. Santanello, S. Kumar, and C. Peters-Lidard, "Representation of soil moisture feedbacks during drought in NASA unified WRF (NU-WRF)," *Journal of Hydrometeorology*, vol. 14, pp. 360-367, 2013. [Online]. Available: <https://doi.org/10.1175/JHM-D-12-069.1>. Accessed: Dec. 22, 2022.
- [17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, 2012, pp. 1097-1105. Accessed: Dec. 22, 2022.
- [18] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780. doi: 10.1162/neco.1997.9.8.1735. Accessed: Dec. 22, 2022.
- [19] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *International Conference on Learning Representations (ICLR)*, 2015. [Online]. Available: <https://arxiv.org/abs/1412.6980> (Accessed: 22/12/2022).
- [20] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint arXiv:1609.04747*, 2016. [Online]. Available: <https://arxiv.org/abs/1609.04747> (Accessed: 22/12/2022).
- [21] T. Tieleman and G. Hinton, "RMSprop: Divide the gradient by a running average of its recent magnitude," *COURSERA: Neural Networks for Machine Learning*, 2012. [Online]. Available: http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_1_ec6.pdf (Accessed: 22/12/2022).