Intelligent Computing Based Forecasting of Deforestation Using Fire Alerts: A Deep Learning Approach

Muhammad Ali Jamshed[a,*], Charalambos Theodorou[b], Tahera Kalsoom[c], Nadeem Anjum[d], Qammer H.Abbasi[a], Masood Ur-Rehman[a,*]

[a] *James Watt School of Engineering, University of Glasgow, UK*
[b] *School of Computer Science and Technology, University of Bedforshire, UK*
[c] *School of Computing, Engineering and Physical Sciences, University of the West of Scotland, UK*
[d] *Department of Software Engineering, Capital University of Science and Technology, Islamabad, Pakistan*

Abstract

Deforestation is depletion of the forest cover and degradation in forest quality mainly through repeated fires, over-exploitation, and diseases. In a forest ecosystem, occurrence of wildfires is a natural phenomena. The curse of global warming and man-made interventions have made the wildfires increasingly extreme and widespread. Though, extremely challenging due to rapidly changing climate, accurate prediction of these fire events can significantly improve forestation worldwide. In this paper, we have addressed this issue by proposing a deep learning (DL) framework using long short term memory (LSTM) model. The proposed mechanism accurately forecasts weekly fire alerts and associated burnt area (ha) utilising historical fire data provided by GLOBAL FOREST WATCH. Pakistan is taken as a case study since its deforestation rate is among the highest in the world while having one of the lowest forest covers. Number of epochs, dense layers, hidden layers and hidden layer units are varied to optimize the model for high estimation accuracy and low root mean square error (RMSE).

*Corresponding author

Email addresses: muhammadali.jamshed@glasgow.ac.uk (Muhammad Ali Jamshed), pampos20101995@hotmail.com (Charalambos Theodorou), tahera.kalsoom@uws.ac.uk (Tahera Kalsoom), nadeem.anjum@cust.edu.pk (Nadeem Anjum), qammer.abbasi@glasgow.ac.uk (Qammer H.Abbasi), masood.urrehman@glasgow.ac.uk (Masood Ur-Rehman)

Simulation results show that the proposed method can predict the forest fire occurrences with 95% accuracy by employing a suitable hyperparameter tuning.

Keywords: Forecasting, estimation, deforestation, deep learning, long short term memory (LSTM), forest fires.

1. Introduction

Forests are instrumental in maintaining the balance between different ecosystems by providing habitat, supplying oxygen, protecting watersheds and acting as a carbon sink. Despite their essentiality, forests are increasingly being destroyed [1]. This forest loss is a combination of deforestation caused by woodcutting, ranching, mining, oil extraction, dam construction, infrastructure development and forest degradation induced by climate changes, wildfires, agriculture, commercial plantation and tree diseases [2]. It was estimated in 2020 that tropics lose trees almost equal to 30 soccer fields every minute. Some of the daunting examples include the Amazon that has lost 17% of its forests in the last 50 years due to human activities [3], the island of Sumatra in Indonesia that has lost 85% of its forests due to oil extraction and pulp and palm plantations [4], Ivory Coast that has lost 26% of its forests in the last 18 years due to cultivation of cocoa [5] and Russia where wildfires have destroyed a record 18.13 Mha of forests in 2021 alone [6]. Pakistan is identified as having one of the fastest deforestation rates and has one of the lowest forest covers in the world [7]. Urgent measures to counter deforestation are necessary to avoid a fast-looming calamity there. Billion Tree Tsunami and Clean Green Pakistan are the two major initiatives taken by the Government of Pakistan recently to slow down the deforestation in the country [8].

A plethora of efforts are being made to accurately map the deforested areas by using satellite images. Use of deep learning (DL) techniques have proven to be effective in estimating the deforestation rates, which in return helps to produce accurate forest maps and enables the relevant authorities to take precautionary measures to combat deforestation [9]. Maretto *et. al* have proposed a DL approach to map the deforested area in Brazilian Amazon through PRODES maps [10]. A DL based method has been employed to classify crops using images from Landsat-8 and Sentinel-1A RS satellites in [11]. A U-Net model to detect the deforested area has been proposed by Isaienkov *et. al* that makes use of the data-set created on Sentinel-2 imagery [12]. Franke *et. al* have developed a spectral mixture based forest disturbance monitoring technique in [13]. This method relies on satellite images to predict tree logging and wildfire activities.

Most of the work available in the literature banks on satellite images to accurately map the deforested areas. Recent advances in technology and ease of availability of historical data facilitates us towards accurate prediction of future wildfires. Being one of the major drivers of deforestation, wildfire prediction can help in the reduction of the deforested areas around the globe. To the best of our knowledge, no structured effort has been made to forecast the wildfire occurrence for the preservation of forests in open literature. This paper attempts to fulfill this gap by proposing an efficient wildfire prediction model taking Pakistan as a case study. The GLOBAL FOREST WATCH platform is used to acquire the historical forest fire data for the country [14]. A DL framework based on long short-term memory network (LSTM) model is used for prediction of weekly fire events and estimated burnt area. The model has appeared to be highly capable of precise forest fire estimations yielding an accuracy level of 95%. It is therefore, plausible to be used as an efficient analytical tool for wildfire trend assessments and provide basis for devising policies and enacting intervention techniques to protect the forest cover.

2. Deep Learning Forecasting

The DL methods have shown good performance in terms of forecasting nonlinear data-sets and solving various time series prediction problems [9]. A number of DL methods are being proposed in literature to forecast the nonlinear data-sets. Some of the popular DL methods include LSTM, deep belief network

(DBN), gated recurrent unit (GRU), artificial neural network (ANN), convolutional neural network (CNN), and support-vector machine (SVM) [15]. A comparison of different DL techniques in [16] has brought out LSTM as a clear winner in dealing with nonlinear data-sets and efficiently making time series predictions. Its ability to remember patterns for short as well as long periods results in better accuracy and help in overcoming the problem of time series prediction. Since, the forest fire data-set obtained from GLOBAL FOREST WATCH platform [14] is strictly non-linear, use of this historical time stamp data to make future time stamp based predictions highly accurate, lead us to the adoption of LSTM in this proposed model.

The LSTM was developed to manage long-term dependencies and for addressing time series prediction problems [17]. The LSTM is a modified version of recurrent neural network (RNN) and is suitable to solve a wide variety of problems [15]. The LSTM is able to include or exclude data from a state of a cell. The structures used are known as gates and are responsible for controlling the data. The LSTM solves the issue of recollection of data over a given period by representing the memory cells and gate units in the form of a neural network. Each memory cell has a state that stores the latest data. Whenever, a memory cell receives new information at a given time instant, the outcome is controlled by the cell state and is updated accordingly. Additionally, if the memory cell receives any other information, it will utilize both the newly received information and the refreshed state information in forecasting [17].

The LSTM based model consists of multiple gates, such as, update gate ξ_u (which is a combination of input gate and forget gate), output gate ξ_o, and forget gate ξ_f, which are used to pass data between cells. Gates combine an activation function with a point-by-point multiplication process. If a number that enters the input gate is 0, then nothing should pass, while 1 allows everything to pass. The LSTM generates the following outputs:

$$\text{Outputs} = \begin{cases} c^{<\bar{t}>} = \xi_u \times c^{N<\bar{t}>} + \xi_f \times c^{<\bar{t}-1>} \\ a^{<\bar{t}>} = \xi_o \times c^{<\bar{t}>}, \end{cases} \tag{1}$$

here, $c^{<\bar{t}>}$ and $a^{<\bar{t}>}$ represent the candidate and the activation values and controls the memory of the LSTM model at a given time stamp $\bar{t}$. The ξ_u, ξ_f and ξ_o are defined as:

$$\xi_u = \delta(W_u[h^{<\bar{t}-1>}, x^{\bar{t}}] + b_u), \tag{2}$$

$$\xi_f = \delta(W_f[h^{<\bar{t}-1>}, x^{\bar{t}}] + b_f), \tag{3}$$

$$\xi_o = \delta(W_o[h^{<\bar{t}-1>}, x^{\bar{t}}] + b_o), \tag{4}$$

here, δ represents the activation function, W_u, W_f, W_o are the weights, and b_u, b_f, b_o are the connection biasing values of ξ_u, ξ_f and ξ_o, respectively. Whereas, $h^{<\bar{t}-1>}$ is hidden state at time instant $\bar{t}-1$ and $x^{\bar{t}}$ is the input at time instant t.

3. Methodology

The LSTM based DL framework is proposed due to the presence of non-linearity in the data-sets. Initially, the pre-processing of the historical forest fire and burnt area (ha) data-sets was carried out followed by the feature selection process. The data-sets were split into two parts. First part was used to train the LSTM model, whereas, the second part was used for testing purposes. Finally, the adjustment of hyperparameters is carried out to minimize the root mean square error (RMSE). A simplified version of the whole process is illustrated in Fig. 1. Moreover, the detailed description of the data-sets and the DL based framework used to forecast the future forest fires is provided in the following subsections.

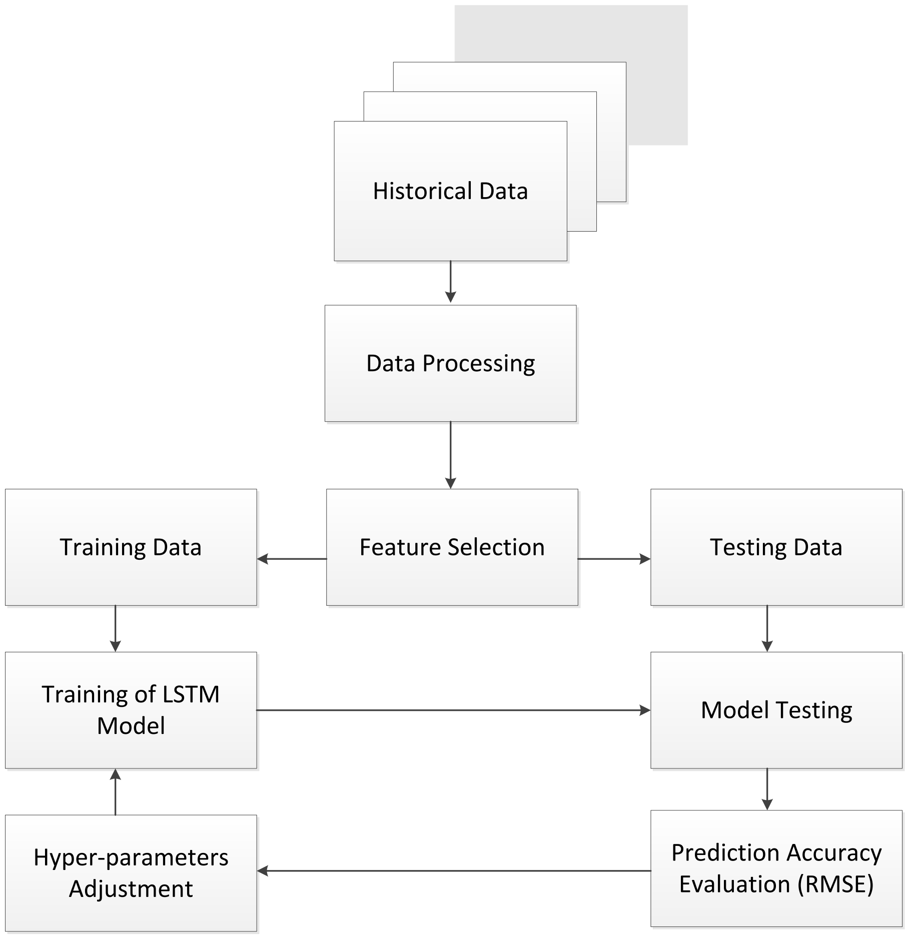

Figure 1: An overview of LSTM training and testing methodology.

3.1. Pre-processing, Training, & Testing

The processing of data-sets was performed using Google colaboratory and Python, in conjunction with a graphics processing unit (GPU). Firstly, the wildfire and burnt area (ha) data-sets were pre-processed by employing a min-max based feature scaling. The min-max function enables the scaling of the features within a specified range. Each feature was scaled and translated individually so that it lies within the range displayed on the training set. Secondly, the processed data-set was divided into training data and testing data. The training

Table 1: Performance comparison between different hidden layers used in the LSTM model

Epochs	3 hidden layers error	Time(minutes)	6 hidden layers error	Time(minutes)
25	0.0007155	7	0.0009452	15
50	0.0004761	14	0.0005676	40
100	0.0004423	30	0.0005541	81
150	0.0004228	45	0.0005428	120

data was used to train the model, while the testing data was used to test the predictions. Pertaining to the availability of a relatively smaller data-sets, i.e., 1450 columns of numeric values, 85% of the data was utilized for training, and 15% data was used for testing. A higher amount of the data used to train the model also helps to reduce discrepancies in predictions. The training data was then passed through LSTM model with different tuning parameters. Finally, the predictions were compared with the testing data to evaluate the accuracy of the forecasting.

3.2. Selection of Hyperparameters

The performance of DL models heavily relies on the selection and tuning of appropriate hyperparameters, as it helps in minimizing the predefined loss function. Inappropriate selection of hyperparameters will effect the accuracy of DL models by increasing the RMSE. Based on the available data-sets, the following hyperparameters were chosen:

- Number of epochs.
- Number of hidden layers.
- Learning rate (lr).
- Number of dense layers.
- Activation function (rectified linear unit (ReLU)).

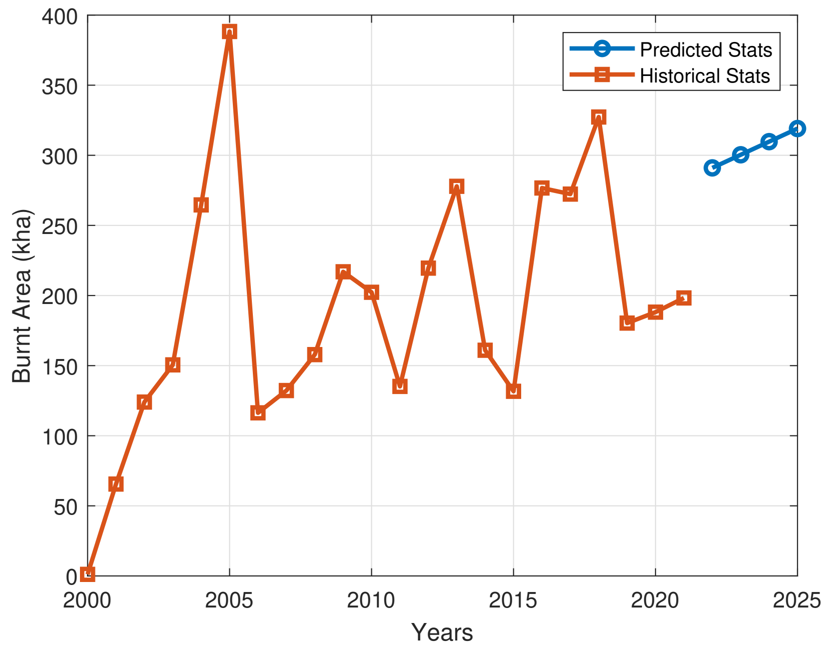

Figure 2: Predicted burnt area in kha from 2022-2025.

- Dropout.

The number of epochs represents the number of iterations the model uses to train itself. In this study, the model was trained for 150 epochs. Table 1 shows a comparison of the LSTM model with 3 and 6 hidden layers. The RMSE with 6 hidden layers appears to be higher. Therefore, the number of hidden layers for the proposed model was set to 3. The value of lr is set to be 0.01 as this value should be kept as lowest as possible. By using a single dense layer, the model achieved a RMSE of 0.00004142 while it garnered a RMSE of 0.00004051 with 2 dense layers. This lead to the selection of 2 dense layers. The over-fitting of data was avoided by fixing the value of dropout to 0.05. The ReLU function showed better performance compared to Sigmoid function and was therefore, selected as the activation function for the proposed LSTM model.

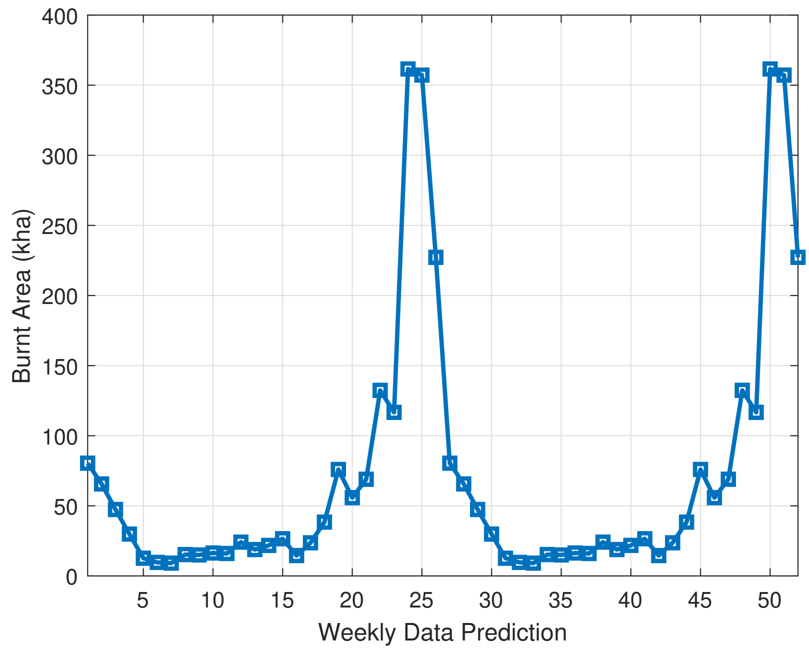

Figure 3: Weekly burnt area prediction in kha for year 2024.

4. Performance Evaluation

The forest fire data for Pakistan covering a period from 1^{st}-January-2012 $\rightarrow$ 1^{st}-October-2021 and the burnt area (ha) data covering a period from 1^{st}-January-2000 $\rightarrow$ 1^{st}-October-2021 was utilized in this work. The input data was in numerical format and consisted of yearly alerts, weekly alerts, burnt area (ha) and total alerts in a given year. Based on pre-defined parameters discussed in Section III, the LSTM model was first trained and then the predictions were generated. The accuracy of the model was evaluated using RMSE, which is a well-known accuracy metric for such models.

Fig. 2 shows the model's predicted reduction in the forest area. A linear increase in burnt area (ha) can be observed for 2022 till 2025. It is estimated that in 2025, the forest area in Pakistan will reduce by 320 kha due to the wildfires. Fig. 4 shows the prediction of weekly fire events for 2023. An estimated increase

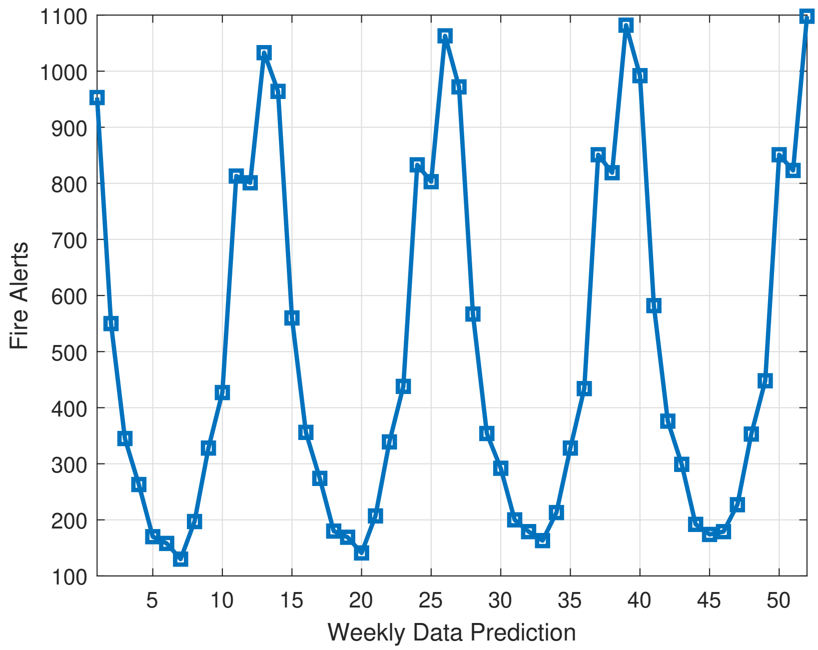

Figure 4: Weekly fire alert prediction for year 2023.

in the weekly fire events from week 15 to week 30 and from week 42 until week 49 can be observed. Fig. 3 exhibits the predicted reduction in forest area on weekly bases for 2024. The proposed model forecasts a high loss in forest area during week 25 and week 52 due to fire events. Based on these predictions, special precautionary measures can be taken to overcome forest fires and hence, reduction in forest areas due to these fire events.

Following the scaling, training, and testing of data-sets, we observed that the values of RMSE strictly depend on the number of layers, units, and epochs used to make predictions. Also, a careful selection of the number of epochs is necessary during the training phase based on the amount of available data to prevent an increase in RMSE. Moreover, it appears that the testing accuracy increases and the training time decreases as the number of dense layers and hidden neurons decreases in the LSTM model. The number of neurons in a hidden layer should not exceed the number of inputs per epoch for training the model to achieve a better accuracy. In the dense layer, the result were also passed through a nonlinear activation function. Therefore, it is important to choose the appropriate number of dense layers for training a model. Proper tuning of the hidden layers also play a crucial role when designing a DL enabled method.

5. Conclusion

In this paper, we have proposed a robust DL based LSTM framework for precise forecasting of the deforestation area using historic forest fire data. Using data from an independent online forum, the future forest fire events in Pakistan as a test case has successfully been predicted with a 95% accuracy. The accuracy of the model has shown a strong reliance on the tuning of hyperparameters. The ability of the proposed model for accurate predictions of the weekly fire events can help the authorities to prepare well ahead and act appropriately to overcome possible future wildfires avoiding consequential deforestation and climate degradation. Targeted future expansion of this work include prediction of forest fire events along with location identification to further aid the forest preservation efforts.

Authors' contributions

Muhammad Ali Jamshed: Conception and design of study, Simulation, Acquisition of data, Writing - original draft. **Charalambos Theodorou:** Conception and design of study, Simulation, Acquisition of data. **Tahera Kalsoom:** Analysis and/or interpretation of data, Writing - original draft. **Nadeem Anjum:** Writing- review and editing.**Qammer H. Abbasi:** Writing-review and editing. **Masood Ur-Rehman:** Conception and design of study, Writing - review and editing.

Competing Interest statement

The authors declare that they have no known competing finan-cial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work has been carried out under the development of ICRG PAK-UK Education Gateway (2020) funded Project No. 310366 - Deforestation in Pakistan: Combating through Wireless Sensor Networks (DePWiSeN).

References

[1] L. V. Gatti, L. S. Basso, J. B. Miller, M. Gloor, L. Gatti Domingues, H. L. Cassol, G. Tejada, L. E. Aragão, C. Nobre, W. Peters, et al., Amazonia as a carbon source linked to deforestation and climate change, Nature 595 (7867) (2021) 388–393.

[2] N. Hosonuma, M. Herold, V. De Sy, R. S. De Fries, M. Brockhaus, L. Verchot, A. Angelsen, E. Romijn, An assessment of deforestation and forest degradation drivers in developing countries, Environmental Research Letters 7 (4) (2012) 044009.

[3] DEFORESTATION AND FOREST DEGRADATION, `https://www.worldwildlife.org`, accessed: 2021-11-17.

[4] E. S. Mcilvride, What messaging factors about the impact of unsustainable palm oil affect commitment to relevant behaviour change?

[5] I.-O. Abu, Z. Szantoi, A. Brink, M. Robuchon, M. Thiel, Detecting cocoa plantations in cote d'ivoire and ghana and their implications on protected areas, Ecological Indicators 129.

[6] Record breaking fires in Siberia, `https://www.greenpeace.org/`, accessed: 2021-12-8.

[7] S. Ullah, T. Gang, T. Rauf, F. Sikandar, J. Q. Liu, R. S. Noor, Identifying the socio-economic factors of deforestation and degradation: a case study in gilgit baltistan, pakistan, GeoJournal (2020) 1–14.

[8] M. Sabir, Y. Ali, I. Khan, A. Salman, Plants species selection for afforestation: A case study of the billion tree tsunami project of pakistan, Journal of Sustainable Forestry (2020) 1–13.

[9] M. Mohammadi, A. Al-Fuqaha, S. Sorour, M. Guizani, Deep learning for IoT big data and streaming analytics: A survey, IEEE Communications Surveys & Tutorials 20 (4) (2018) 2923–2960.

[10] R. V. Maretto, L. M. Fonseca, N. Jacobs, T. S. Körting, H. N. Bendini, L. L. Parente, Spatio-temporal deep learning approach to map deforestation in amazon rainforest, IEEE Geoscience and Remote Sensing Letters 18 (5) (2020) 771–775.

[11] N. Kussul, M. Lavreniuk, S. Skakun, A. Shelestov, Deep learning classification of land cover and crop types using remote sensing data, IEEE Geoscience and Remote Sensing Letters 14 (5) (2017) 778–782.

[12] K. Isaienkov, M. Yushchuk, V. Khramtsov, O. Seliverstov, Deep learning for regular change detection in ukrainian forest ecosystem with sentinel-2, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14 (2020) 364–376.

[13] J. Franke, P. Navratil, V. Keuck, K. Peterson, F. Siegert, Monitoring fire and selective logging activities in tropical peat swamp forests, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 5 (6) (2012) 1811–1820.

[14] M. C. Hansen, P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, et al., High-resolution global maps of 21st-century forest cover change, science 342 (6160) (2013) 850–853.

[15] S. Pouyanfar, S. Sadiq, Y. Yan, H. Tian, Y. Tao, M. P. Reyes, M.-L. Shyu, S.-C. Chen, S. S. Iyengar, A survey on deep learning: Algorithms, techniques, and applications, ACM Computing Surveys (CSUR) 51 (5) (2018) 1–36.

[16] K. Amarasinghe, D. L. Marino, M. Manic, Deep neural networks for energy load forecasting, in: 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE), IEEE, 2017, pp. 1483–1488.

[17] S. Hochreiter, J. Schmidhuber, Lstm can solve hard long time lag problems, Advances in neural information processing systems (1997) 473–479.