Forest Fire Hotspot Prediction Model Using GAF CNN-LSTM

Akmal Raafid Taufiqurrahman
Fakultas Informatika
Telkom University
Bandung, Indonesia
akmalraafid@students.telkomuniversity.ac.id

Irma Palupi
Fakultas Informatika
Telkom University
Bandung, Indonesia
i*rmapalupi@telkomuniversity.ac.id*

***Abstract* — Forest and land fires has been one of the problematics in Indonesia, particularly in Borneo. Indonesia contributes for a considerable proportion of Southeast Asia's forest area and currently occurring one of the world's highest deforestation rates, second to Brazil. Forest fires occur periodically during the dry season where the lands are covered by peat land. Along with climate change, regions that have a great fuel consumption are becoming suspectible to the intensity of forest fire. Forest fire prediction has become important to prevent forest fire from occurring and to make first response when the fire occurs. The model for forest fire prediction represents an essential tool to predict forest fire risk, damage, forest fire monitoring and extinction phase, and to assist in the fire control planning and protecting both human life and property. We proposed a combination of three methods called GAF-CNN-LSTM that uses image input to predict carbon emission as the factor of forest fire. This study aims to create a model that can predict forest fire hotspot accurately. The result shows that the proposed method performed better than the LSTM only method by having a better loss value and loss reduction in each iterations.**

Keywords— Forest Fire, Prediction

I. INTRODUCTION

Forest and land fires has been one of the problematics in Indonesia, particularly in Borneo. Indonesia contributes for a considerable proportion of Southeast Asia's forest area and currently occurring one of the world's highest deforestation rates, second to Brazil [1]. Forest fire is a part of nature caused by natural (lightning) or unnatural (human error) causes. It occurs periodically during the dry season where the lands are covered by peat land. It means that the land covered may determine the level of fire event. The vegetation type and soil characteristics also contribute to the development of fire occurred in forest fire. Along with climate change, regions that have a great fuel consumption are becoming suspectible to the intensity of forest fire. Forest fire produces an abundant amount of greenhouse gases to the atmosphere and becomes one of the causes of the interannual variability in the growth rates of several trace gases. This may cause public health problems since

pollutant emissions from large-scale forest fire include carbon monoxide (CO). According to the data from Ministry of Environment and Forestry, in 2019 fire burnt approximately 1.6 million ha of forest and land in Indonesia where the most affected regions were Central, West and South Borneo.

Forest fire prediction has become important to prevent forest fire from occurring and to make first response when the fire occurs [2]. If fires are not detected early, they may become out of control and the consequences are often disastrous. The model for forest fire prediction represents an essential tool to predict forest fire risk, damage, forest fire monitoring and extinction phase, and to assist in the fire control planning and protecting both human life and property. Today many fire risk models make use of forest fire databases to construct and assess probabilistic models.

Many recognition-based problems commonly can be solved with classical algorithms such as decision tree (DT) and random forest (RF). However, these classical algorithms all require complex and time-consuming feature engineering, which requires not only designing of extracted features manually, but also features selection or dimensionality reduction to screen out the best representative features [3]. The advancements in Machine Learning, called Deep Learning emerged as machine learning algorithms. Amongst contemporary approaches, Deep Learning.

(DS) has risen into an attractive approach in this topic. Deep Neural Networks provide a potent mechanism that allows for learning complex mappings from raw data automatically, avoiding the need for developing hand-crafted features [4]. This algorithm shows an exceptional performance for computer vision application, including object detection and image classification. These advancements inspired the framework to encode time series data as distinct types of images, namely Grammian Angular Field (GAF). This enables the use of computer vision for classification. Using coordinate system, GAF images are represented as matrix where each element is trigonometric sum between different time intervals. Another method we are using is called Convolutional Neural Networks (CNN). CNN converts transitional invariance within structures by extracting features from image input. The features are then inserted into the Long Short-term Memory (LSTM) to classify weight, posture and then updated through repeated training process until generates the best performance. Compared to the classical algorithms, CNN got better recognition performance and training time.

This study is limited to only using multivariate timeseries data as the input data and only researches the forest fire in Borneo. Based on the previous research, the forest fire hotspot prediction model will be implemented using GAF CNN-LSTM by processing time-series data. The datasets used in this study will be taken from public sources which are daily data from 1998 to 2022. The data obtained will be converted to an image using GAF method and CNN method will be implemented to classify multivariate datasets while the LSTM will be implemented to predict the hotspot of the forest fire by considering the given variables. This study aims to find a model that can predict forest fire, find the related variables that supports the occurrence of forest fire and implement the suitable model for the given problems and evaluate the result to compare.

II. LITERATURE STUDY

A. Forest Fire

Forest fire is one of the most dangerous disasters in the world. Forest fires cause inestimable loss of forest resources, property and threaten people's lives. Along with climate change, many forests are threatened by forest fire, especially Indonesia's Borneo. Indonesia contributes for a considerable proportion of Southeast Asia's forest area and currently occurring one of the world's highest deforestation rates, second to Brazil [1].

Forest fires occurrence can be affected by vegetation, slope, altitude, humidity and temperature [5]. Other than natural causes, forest fire can be caused by human activity. Intentional forest fires by humans occur to open new land and are started by local farmers, private companies and government agencies to clear the land before panting crops. Forest fires in Indonesia are closely related to the climate disasters Moreover, forest fire contributes much carbon emissions into the atmosphere and could double if not stopped. carbon emissions from forest fires in Indonesia were estimated to be almost double compared to forest fires in the Amazon, Brazil.

Forest fire prediction has become highly important to predict forest fire risk, back up forest fire monitoring and extinction, and assist in fire control planning and resource management to reduce casualty in property and prevent human casualty.

B. Gramian Angular Field

Gramian Angular Field (GAF) is a data transformation method that converts time series data into a 2D texture images. GAF maps time series data into an image and retains the original time series features and the features will be learned by CNN afterwards [6,7,8].

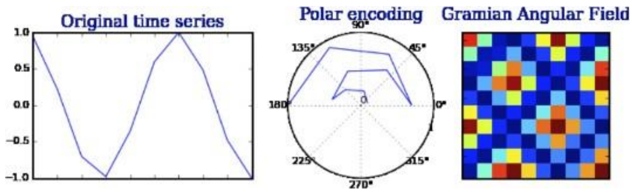


Figure 1. Converting Multivariate Time Series Input using GAF

In this problem, we can represent the images converted from the input as a matrix where each matrix contains information regarding the variables used in the input data. Each images represents the correlation between each pair of values from time series. The data inputted is converted into a grid size and then the image features can be identified by the color such as black for value one and white for value zero. The color grading in the image means that the value is ranging from zero to one or between one and zero. GAF converts a time series data from cartesian coordinates into polar coordinates system by using formulas (1- 2). Let X be a time series data consists of n time indeces, which the values are in range $[-1,1]$. For each xi, the value is transformed to polar coordinate using arccos function as given by formula (1), and then the field matrix is computed by formula (2) represents the interaction the value at a time index to other.

$$X = \{x_1, x_2, \dots, x_n\} \begin{cases} \emptyset_i = \arccos(x_i), -1 \le x_i \le 1, x_i \in X \\ r = \frac{t_i}{n}, t_i \in [1, n] \subset \mathbb{Z} \end{cases} \quad (1)$$

$$GAF_{matrix} \begin{Bmatrix} \cos(\emptyset_1 + \emptyset_1) \cos(\emptyset_1 + \emptyset_2) \dots \cos(\emptyset_1 + \emptyset_n) \\ \cos(\emptyset_2 + \emptyset_1) \cos(\emptyset_2 + \emptyset_2) \dots \cos(\emptyset_2 + \emptyset_n) \\ \vdots \\ \cos(\emptyset_n + \emptyset_1) \cos(\emptyset_n + \emptyset_2) \dots \cos(\emptyset_n + \emptyset_n) \end{Bmatrix} \quad (2)$$

The polar coordinate system uses the length and the degree of tilt as the measurement of a certain position. As shown in Fig 2.2, the graphs are bended towards the relative calculated angle as the time increases. GAF normalizes the input to [-1,1] interval that should be beneficial in a classification problem.

C. Convolutional Neural Network

Convolutional Neural Network (CNN) is a feature extraction method that is usually used in image recognition. CNN is basically used for time-series variables significance inspection by examining the influence of each channel on the weights of the output layer contributing to the predicted outcome [4,3]. The images produced in GAF are then analyzed to learn certain patterns to help the model predictions. The color gradients are used to highlight the state of the tree in the forest, whether it is on fire, not on fire or empty land. By extracting those features, the images are then can be converted into a set of values of states and variables that affect the state.

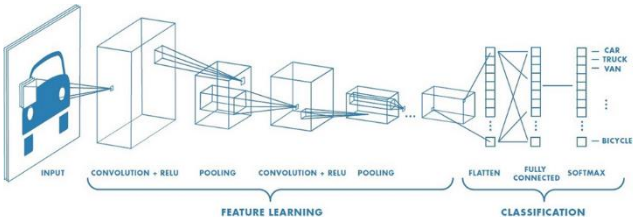


FIGURE 2. Convolutional Neural Network Structure

The input images have a k-dimension since the data converted is a time series data. The image converted by GAF is needed to go through convolution layer, pooling layer, fully connected layer. In the convolution layer, the image dimensions are going through a filter that determines how large the dimension of the data is and how many pixels are considered in the calculations. In the pooling layer, similar process is applicated through the data as happened in the convolutional layer, the difference is that we need to decide whether the average or maximum value are taken from the data to preserves a small feature in a smaller pixel that are crucial.

D. Long Short-term Memory Network

Long Short-term Memory Network (LSTM) is a recurrent neural network (RNN) architecture that memorize values over intervals. LSTM address RNNs problem by segmentation gating functions as their state dynamics. It composed of cell, input gate and forget gate. The cell is responsible for memorizing values over time intervals and the gates regulate the flow of information into and out of the cell. Hence, LSTM is often used for analyzing sequences such as videos, images or multidimensional environmental [9].

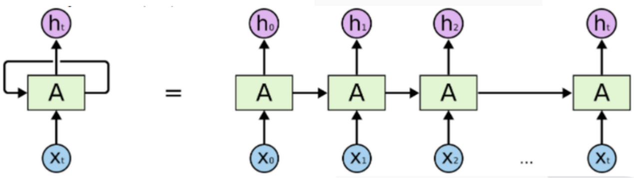


FIGURE 3. Structure of RNN

RNN consists of loops which allow information to be kept on running. LSTM maintains a hidden vector and a memory vector, which controls state update and outputs [10]. The RNN architecture is able to represent the past output values from current output calculations to produce an ideal for autocorrelation structures from time series data.

E. Mean Absolute Error

Mean Absolute Error (MAE) is a measure of error for measuring between paired observations reprsenting the same phenomenon. MAE is a common error measure of forecasting errors in time series analysis.

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \qquad (3)$$

III. SYSTEM DESIGN

We proposed a hotspot prediction system that processes multivariate time series data gathered from ERA5 and GFED. The data itself contains variables such as carbon emissions as the target value, latitude, and longitude. The input data contains temperature, precipitation, wind speed, dewpoint temperature, and will be processed through GAF to be converted into an image. Then, the image's features will be extracted afterwards by CNN to be classified and then, the LSTM will provide further prediction based on the data given. The variables will also be analyzed for the influence of forest fire since forest fire will occur depending on some external factors contributing to the abundance of external factors.

A. Flow Diagram
B.

To create the model for forest fire hotspot prediction using GAF-CNN-LSTM, the steps as shown in figure 4 will be implemente

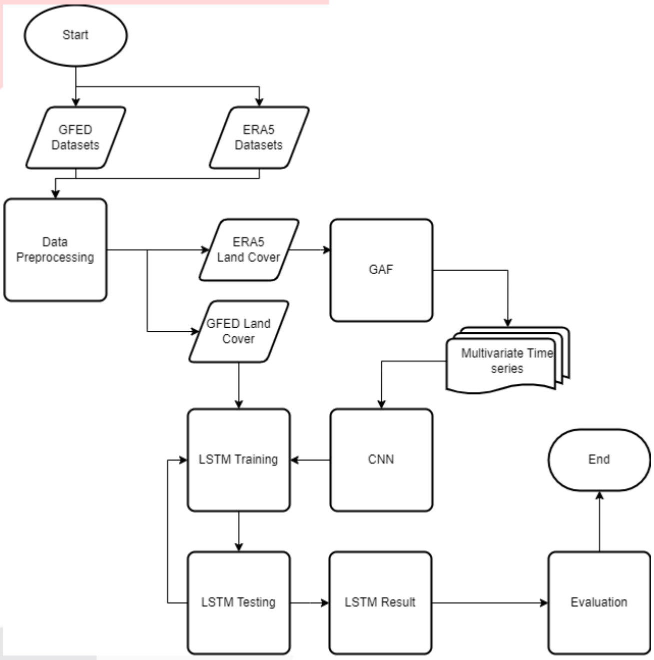


FIGURE 4. System Design of Forest Fire Hotspot Prediction Model

Emission data obtained are in format of yearly and will be merged into 24 years with monthly format and to be processed as the target value for this model in format of monthly. While the climate data needs to be pre-processed to eliminate noise and to reduce the number of errors by filtering the data into land cover only, since the data obtained also covers the sea and land. After filtering, data will be converted into images for each variables using GAF, then the resulted image's features will be extracted using CNN, the extracted features will be learned by LSTM as the X train and the emission data will be trained as the Y train to create a prediction based on the features. The input conversion can be detailed into several steps visualized in Figure 5

FIGURE 5. Input Conversion using GAF

The time series coordinates data will be converted into polar coordinate system by GAF encoding. The grid represents coordinates of the forest. After being converted, the image features will be learned using CNN-LSTM to create accurate forecasting in multivariate time series. Each variable in data will be transformed into 4x4 matrix, all four variables matrices will be concatenated into a new 8x8 matrix and will be also added a matrix of time, longitude and latitude so the final matrix will be in the form of 8x16.

C. Exploratory Data Analysis

The dataset is collected from Global Fire Emissions Database (GFED), which refers to the forest fire data history [10]. The dataset used is a dataset of GFED from 1998 to 2022. The dataset contains 4 variables, which are emissions, time, latitude, and longitude. In this case, emissions mean the carbon emissions that are calculated in a certain area. Longitude is ranging from 108.5 as the western most longitude to 119.5 as the eastern most longitude and latitude ranging from -4.25 as the northernmost latitude to 6.25 as the southernmost latitude. The dataset contains 541800 rows with 4 columns and has a 0.25-degree latitude by 0.25-degree longitude spatial resolution.

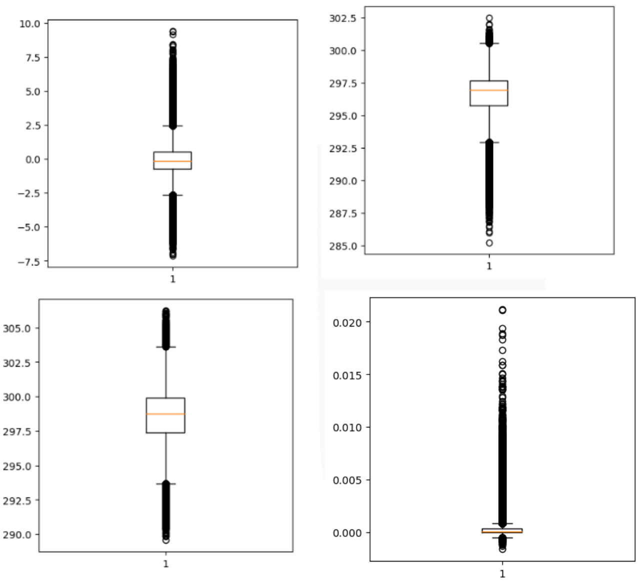


FIGURE 6. Boxplot of four climate variables

As seen on Figure 6, each variable has the distribution of data near maximum and minimum value. It can be said that the data is distributed around the maximum and minimum value

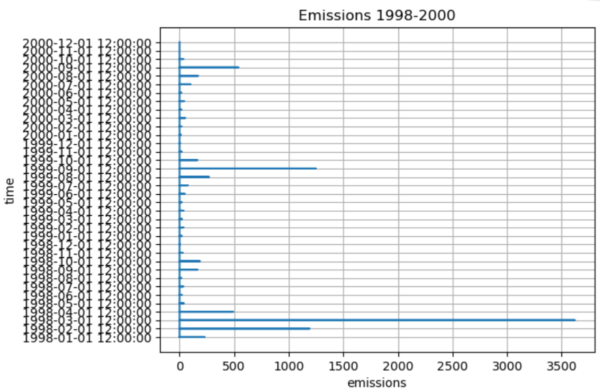


FIGURE 7. Emissions From 1998-2000

From Figure 7, the cycle of emission increase occurs in each year within the 9th month and the 3rd month. It can be stated that emissions production in both months are higher due to the usage of fuel, deforestation, and forest fire. From the figure, it can be concluded that emissions can be considered as the factor of forest fire and in this case, it can be used as the target value of forecast.

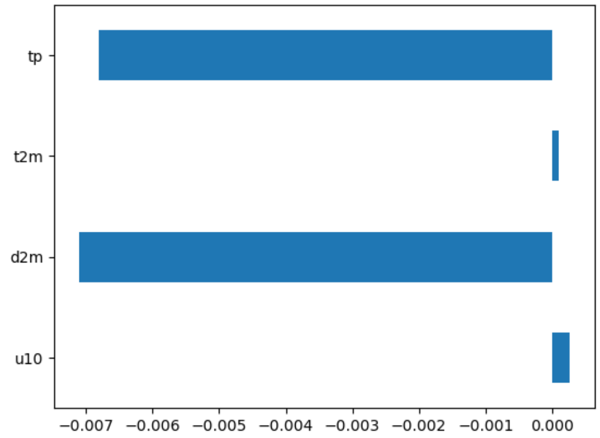


FIGURE 8. Correlation between four factors to emissions

From Figure 8, total precipitation and dewpoint temperature have a strong negative correlation which means that both values can be a factor in reducing the amount of carbon emissions. It means that the higher the value of both factors can also reduces the value of carbon emissions. Whereas the windspeed and 2 meter temperature factor has a weak positive correlation which means that both factors has an insignificant effects on the carbon emissions value.

D. Preprocessing Data and Feature Selection

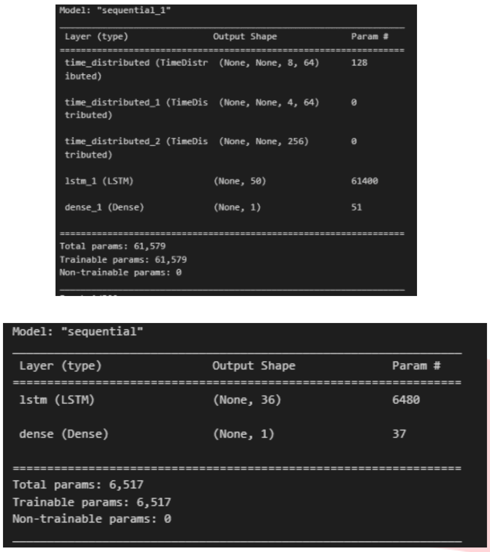
In data preprocessing, the GFED datasets will be merged into timeseries data from 1998 to 2022 in the form of CSV. Then the ERA5 Climate Data will be re-grided to match the GFED latitude and longitude resolution using two-dimensional cubic spline interpolation. After re-grid, ERA5 dataset will be filtered into the land cover data only. In this case, all considered climate factors such as emissions, total precipitation, dewpoint temperature, temperature, and wind component will be used as the training variables and emissions will also be used as the target value for the prediction. All the timestamps are in the format of “yyyy-mm-dd 12:00:00” since the value of the variables are taken at the time of 12:00.

E. CNN-LSTM Implementation

In this section, the 8x8 matrix will be processed to a LSTM model and CNN-LSTM model to be compared. The matrices contain information of longitude, latitude time ,10m Wind Speed, 2m Temperature, 2m Dew point and Total precipitation. All this information will be fitted into the model to an emissions variable as the target variable. The model’s performance will be compared to be evaluated.

In this case, we used the first 21 years of data as the training data and the last 3 years as the testing/validation data. The model will process the data from each location where the location can be represented in every 300 data.

The CNN-LSTM model has a Convolutional layer, Max Pooling layer, and LSTM layer. The LSTM model only uses a LSTM layer.



```
Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 time_distributed (TimeDistr  (None, None, 8, 64)      128
 ibuted)

 time_distributed_1 (TimeDis  (None, None, 4, 64)      0
 tributed)

 time_distributed_2 (TimeDis  (None, None, 256)        0
 tributed)

 lstm_1 (LSTM)               (None, 50)                61400

 dense_1 (Dense)             (None, 1)                 51

=================================================================
Total params: 61,579
Trainable params: 61,579
Non-trainable params: 0
```

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm (LSTM)                 (None, 36)                6480

 dense (Dense)               (None, 1)                 37

=================================================================
Total params: 6,517
Trainable params: 6,517
Non-trainable params: 0
_________________________________________________________________
```

FIGURE 9. Models Architecture

F. Evaluation Method

The model creates predicted value that can be compared into the actual value, where both values are needed to calculate the Mean Absolute Error value. The MAE represents the difference between the predicted value and the actual value where the MAE value can be considered as good if the value is closer to 0. In this case, we want to create a model that has a good performance in solving this problem.

IV. EVALUATION

A. Evaluation Method

By constructing both LSTM and CNN-LSTM model, for predicting the emissions CNN-LSTM has a smaller loss and smaller value loss. The loss value is getting smaller in each iteration for both models. But, the CNN-LSTM has better loss for this form of problem. In this case, we model a program for creating a prediction based on location where in each location we use 21 years data as the training data and the last 3 years as the testing/validation data. The results of performance can be seen in Figure 10

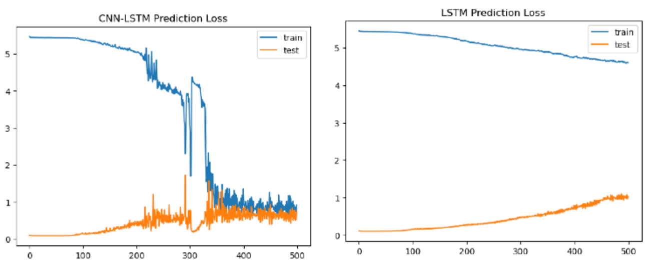


FIGURE 10. The Performance of LSTM and CNN-LSTM

The model has a small mean absolute error value, running the model resulted in 0.41 of mean absolute error value. By looking at the graph, it can be said that the CNN-LSTM model performs better in reading the patterns of the time series value. The model learns the patterns in 500 iterations and could create predictions better than the LSTM model. The LSTM model needs more iterations to learns the patterns. But the result of predictions are still to far than the CNN-LSTM model. It may be effective but not efficient to use since it took more time than the CNN-LSTM model. It can be stated that the CNN-LSTM model is the better option in treating multivariate time series data.

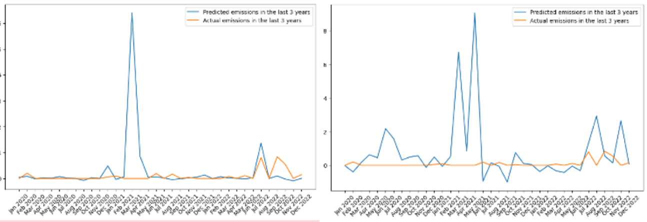


FIGURE 11. The Predicted and Actual Emissions value

CNN-LSTM Predicted emission has a slightly different result than the actual value. But, seen from the MAE value of 0.41, it can be considered as slightly similar to the actual value. As seen in Figure 11, we can see that the CNN-LSTM can learn the pattern for the prediction better than the LSTM model.

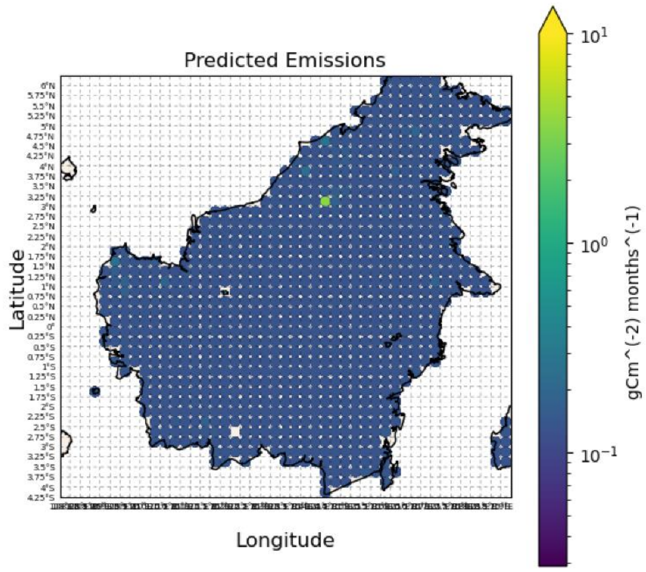


FIGURE 12. The map of predicted emissions in each location in the year of 2022

From the result of the model, we can see that the model predicts the emissions in every location in a certain year it can be seen that there's a point where it has a brighter color than the other, it can be concluded that the location that has a brighter color can represent a location where there is a probability of large emissions and can lead to forest fire.

B. Analysis

Model	Mean Squared Error	Mean Absolute Error
LSTM	5.239902674483781	0.6916610265611519
CNN-LSTM	1.1416031800796107	0.411590413370255574

Table 1. Model's Performance in each Performance Metrics

The performance of CNN-LSTM model is better than the LSTM model in forecasting the time series data.

The model's performances shows that the CNN-LSTM model is better than LSTM model by having performance metric value closer to zero than LSTM model. It concludes that the CNN-LSTM performs better in handling the multivariate time series data

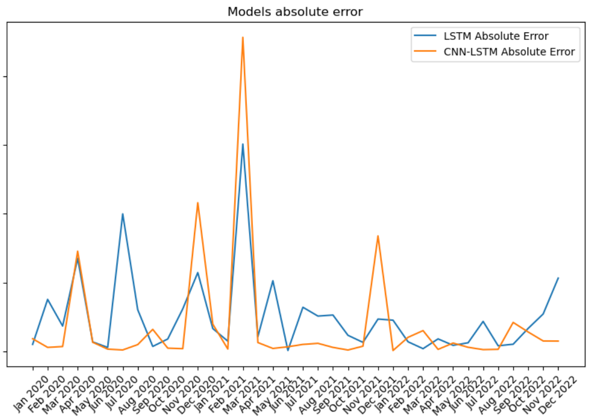


FIGURE 12. Models Absolute Error

Figure 12 states both models absolute error for the last 3 years data in each month. It concludes that the CNN-LSTM has better knowledge in reading the pattern so that in several month the error is closer to zero, but in some cases it can goes to eight since the model haven't learn the pattern. From the figure it can be concluded that in terms of error value the CNN-LSTM is better since it has a smaller error value.

From the result we can see that the CNN-LSTM has no significant outcome where the similarity between actual and predicted value is still slightly far to be considered as similar. Thus, the model can be tuned to get a better performance by changing the parameters values. This can be caused by the preprocessing treatment. For example, the interpolation is done before the data is filtered into land cover only. It can cause loss of value whether it's significant or insignificant.

C. CONCLUSION

The CNN-LSTM model can be a better solution in forecasting the multivariate time series problem. The performance of CNN-LSTM model is better than the LSTM model in forecasting the time series data. The CNN-LSTM model is better than LSTM model by having performance metric value closer to zero than LSTM model. The CNN-LSTM model learns the patterns in 500 iterations and could create predictions better than the LSTM model. The LSTM model needs more iterations to learn the patterns. But the result of LSTM model predictions is still too far than the CNN-LSTM model.

It can be concluded that the CNN-LSTM models' similarity between actual and predicted value is slightly far from the actual value. The CNN-LSTM model still hasn't learned the pattern in several cases. Also, the model's absolute error that in some cases can be greater than one. Although, the result of the forecast can be improved by better hyper parameter tuning, adding more related factor, and choosing the correct preprocessing treatment of the data. For future work others may use the ConvoLSTM as the prediction model to be compared with the CNN-LSTM

REFERENCES

[1] S. Yin, X. Wang, M. Guo, H. Santoso, and H. Guan, "The abnormal change of air quality and air pollutants induced by the forest fire in Sumatra and Borneo in 2015," Atmos Res, vol. 243, Oct. 2020, doi: 10.1016/j.atmosres.2020.105027.

[2] G. E. Sakr, I. H. Elhajj, G. Mitri, and U. C. Wejinya, "Artificial Intelligence for Forest Fire Prediction," 2010.

[3] H. Xu et al., "Human activity recognition based on gramian angular field and deep convolutional neural network," IEEE Access, vol. 8, pp. 199393–199405, 2020, doi: 10.1109/ACCESS.2020.3032699.

[4] F. Viton, M. Elbattah, J. L. Guerin, and G. Dequen, "Heatmaps for Visual Explainability of CNN-Based Predictions for Multivariate Time Series with Application to Healthcare," in 2020 IEEE International Conference on Healthcare Informatics, ICHI 2020, Nov. 2020. doi: 10.1109/ICHI48887.2020.9374393.

[5] L. Si et al., "Study on forest fire danger prediction in plateau mountainous forest area," Natural Hazards Research, vol. 2, no. 1, pp. 25–32, Mar. 2022, doi: 10.1016/J.NHRES.2022.01.002.

[6] Zhang, G., Si, Y., Wang, D., Yang, W., & Sun, Y. (2019). Automated Detection of Myocardial Infarction Using a Gramian Angular Field and Principal Component Analysis Network. IEEE Access, 7, 171570–171583. https://doi.org/10.1109/ACCESS.2019.2955555

[7] Hong, Y. Y., Martinez, J. J. F., & Fajardo, A. C. (2020). Day-Ahead Solar Irradiation Forecasting Utilizing Gramian Angular Field and Convolutional Long Short-Term Memory. IEEE Access, 8, 18741–18753. https://doi.org/10.1109/ACCESS.2020.2967900

[8] Bagnall, A., Dau, H. A., Lines, J., Flynn, M., Large, J., Bostrom, A., Southam, P., & Keogh, E. (2018). The UEA multivariate time series classification archive, 2018. http://arxiv.org/abs/1811.00075

[9] S. Alhirmizy and B. Qader, "Multivariate Time Series Forecasting with LSTM for Madrid, Spain pollution."

[10] F. Karim, S. Majumdar, H. Darabi, and S. Harford, "Multivariate LSTM-FCNs for time series classification," Neural Networks, vol. 116, pp. 237–245, Aug. 2019, doi: 10.1016/j.neunet.2019.04.014.

[11] H. Lee, K. Yang, N. Kim, and C. R. Ahn, "Detecting excessive load-carrying tasks using a deep learning network with a Gramian Angular Field," Autom Constr, vol. 120, Dec. 2020, doi: 10.1016/j.autcon.2020.103390.

[12] T. Oates and Z. Wang, "Encoding Time Series as Images for Visual Inspection and Classification Using Tiled Convolutional Neural Networks." [Online]. Available: www.aaai.org

[13] I. N. Wahyuni, A. Shabrina, and A. L. Latifah, "Investigating Multivariable Factors of the Southern Borneo Forest and Land Fire based on Random Forest Model," in ACM International Conference Proceeding Series, Oct. 2021, pp. 71–75. doi: 10.1145/3489088.3489115.