

Multi-step CNN forecasting for COVID-19 multivariate time-series



Haviluddin ^{a,1,*}, Rayner Alfred ^{b,2}

^a Department of Informatics, Faculty of Engineering, Universitas Mulawarman, East Kalimantan, Indonesia

^b Knowledge Technology Research Group, Faculty of Computing and Informatics, Universiti Malaysia Sabah, Sabah, Malaysia

¹ haviluddin@unmul.ac.id; ² ralfred@ums.edu.my

* corresponding author

ARTICLE INFO

Article history

Received March 4, 2023

Revised April 16, 2023

Accepted April 26, 2023

Available online May 2, 2023

Keywords

Multivariate time-series

CNN

Smoothing technique

COVID-19

ABSTRACT

The new coronavirus (COVID-19) has spread to over 200 countries, with over 36 million confirmed cases as of October 10, 2020. As a result, numerous machine learning models capable of forecasting the epidemic worldwide have been produced. This paper reviews and summarizes the most relevant machine learning forecasting models for COVID-19. The dataset is derived from the world health organization (WHO) COVID-19 dashboard, and it contains official daily counts of COVID-19 cases, fatalities, and vaccination use reported by countries, territories, and regions. We propose various convolutional neural network (CNN) based models such as CNN, single exponential smoothing CNN (S-CNN), moving average CNN (MA-CNN), smoothed moving average CNN (SMA-CNN), and moving average smoothed CNN (MAS-CNN). Here, MAPE and MSE are used to assess the suggested models. MAPE is frequently used to compare accuracy across time series with different scales. MSE, the model must strive for a total forecast equal to the entire demand. That is, optimizing MSE seeks to create a forecast that is right on average and so unbiased. The final result shows that SMA-CNN outperformed its baselines in both MAPE and MSE. The main contribution of this novel forecasting approach is a more accurate result as a base of the strategy of preventing COVID-19 spreads.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



1. Introduction

Past disease outbreaks have affected humanity, making WHO and national authorities worldwide fight these pandemics. The COVID-19 pandemic, first found in Wuhan, China, in December 2019, remains a major problem [1]. According to the WHO, 213 nations and territories have COVID-19 [2]. COVID-19 can spread through direct physical contact, coughing, or sneezing [3]. In this case, COVID-19 has a 14-day incubation period, which is key to its spread [4]. Therefore, the precise projection of recovered and contaminated COVID-19 cases is essential for monitoring the outbreak and implementing effective methods to prevent its spread.

Biomedical informatics is important to any study attempt to treat COVID-19 patients. Artificial intelligence (AI) is becoming more common in healthcare systems to detect diseases [5] and perform clinical diagnoses [6]. This outbreak demonstrates the need and viability of utilizing AI to anticipate outbreaks. AI tools have excellent feature extraction capabilities [7], which help public health experts make decisions. AI can quickly extract information from health reports [8], social media [9], news [10], and media [11]. AI approaches have advanced in the recent decade. Machine learning [12] and deep learning [13] can automatically extract significant elements from complex data. Moreover, deep learning

has advanced computer vision [14], robotics [15], medical imaging [16], chemistry [17], and forecasting [18]. Several studies in the literature focused on modeling, predicting, and forecasting COVID-19 spread based on recorded COVID-19 time series data to comprehend and manage this pandemic.

Deep learning approaches have gained popularity in time-series modeling and analysis due to their generalization and nonlinear approximation [19]. Deep learning models are created by automatically combining neural network layers and extracting significant information from vast data [20]. This approach has been studied in several machine-learning applications. For example, the LSTM model predicts new COVID-19 cases in Canada from January-March 2020 [21]. In Russia, Peru, and Iran, an enhanced LSTM model predicts COVID-19 epidemic patterns [22]. Moreover, SVR, LSTM, BiLSTM, and GRU are used to forecast COVID-19 time-series data in 10 nations, and the result shows that BiLSTM data accessible through June 27, 2020, show higher performance [23].

Due to its performance, deep learning has been effectively applied to various real-world prediction challenges, including time-series forecasting. They make accurate forecasts despite the noisy and chaotic character of time-series forecasting. CNN is a popular, efficient deep learning approach [24]. CNN models can filter input data noise and extract more valuable characteristics for the final prediction model [25]. Standard CNNs are "feed-forward neural networks" that use filters and pooling layers, well-suited for spatial autocorrelation data but not complex and extended temporal relationships. Therefore, removing noisy samples improves temporal data representation and forecasting system accuracy by highlighting relevant patterns. Smoothing strategies help track data seasonality and improve deep learning performance [26]. Due to their simplicity and firm performance in time series forecasting, most strategies use moving averages [27] and seasonal exponential smoothing [28]. Smoothing improves interpretability and integrates the data series changing pattern into the prediction model. Furthermore, a CNN model automatically generates forecasts from the smoothed results.

This paper proposed various CNN-based models such as CNN, single exponential smoothing CNN (S-CNN), moving average CNN (MA-CNN), smoothed moving average CNN (SMA-CNN), and moving average smoothed CNN (MAS-CNN). These multi-step CNN models may provide an accurate multivariate time-series analysis by modifying CNN to select and optimize the best smoothing techniques. The model and the baselines were then implemented in the various time series from the big five countries' COVID-19 dataset. The contribution of the research are.

- To increase the accuracy of the multivariate time-series forecasting analysis by employing innovative CNN in conjunction with smoothing approach optimization.
- To implement an innovative forecasting model that can adapt to different multivariate time series COVID-19 datasets from the United States of America, India, Brazil, France, and Germany.
- To produce a higher level of prediction accuracy by experimental study and validation of the model compared to the CNN, S-CNN, MA-CNN, SMA-CNN, and MAS-CNN.

The remaining parts of this work are structured as described below. The section titled "CNN-based Forecasting Using Smoothing Approach" provides a detailed explanation of the smoothing technique used for the CNN time series and an explanation of the experimental design. The dataset, the data normalization procedure, the forecasting process, and the key performance indicator are all presented in the "Materials and Methods" section. The "Results and Discussion" section presents the findings and an in-depth examination of the experiments. A summary of the research is provided here, along with a discussion of the numerical experiments. The last section, conclusions, gives a summary of the overall findings as well as potential areas for further research.

2. Method

To carry out research in a manner that is more methodical, we planned the experiment in the manner depicted in Fig. 1. We used a variety of datasets to evaluate the effectiveness of several smoothed CNN compared to basic CNN. Fig. 1 shows that the experimental design employed in this investigation included five scenarios. The five scenarios are as follows: (1) data is processed directly with CNN; (2)

data is smoothed using single exponential smoothing and then processed with CNN (S-CNN); (3) data is smoothed using MA and then processed with CNN (MA-CNN); (4) data is smoothed using CNN. Smoothing using single exponential smoothing first, then MA (SMA-CNN); and (5) smoothing using MA first, followed by smoothing using single exponential smoothing again, which is then processed using CNN (MAS-CNN).

The data quality can be improved by using data smoothing [29]. When applied to time-series data, the smoothing method gets excellent results after removing any outliers that could be present in the data [30]. This method can be easily comprehended and applied successfully in the new study without referring to or taking parameters from previous investigations. By taking the average of the previous values in a time series, smoothing processes make predictions more accurate. The algorithm provides a weighting value assigned to past observations to reduce noise, smooth the value of fluctuations in the data being utilized, and anticipate future values. In general, there are various common types of data smoothing. Single exponential smoothing (S) [31] and moving average (MA) [32] are two more frequent types of data smoothing. Smoothing is a strategy that can assist researchers in predicting trends when they are asked to do a forecasting task.

The exponential window function is utilized in the single exponential smoothing (S) method, a rule-of-thumb approach to smoothing time-series data [33]. Exponential functions are used to apply weights diminishing at an exponential rate over time. It is easy to understand and apply when making judgments based on the user's past assumptions, such as seasonality, and it does not require much time. The Moving Average (MA) is a formula used to examine data points [34]. It begins with creating a series of averages of various subsets of the complete dataset and then continues connecting those averages in the shape of a line.

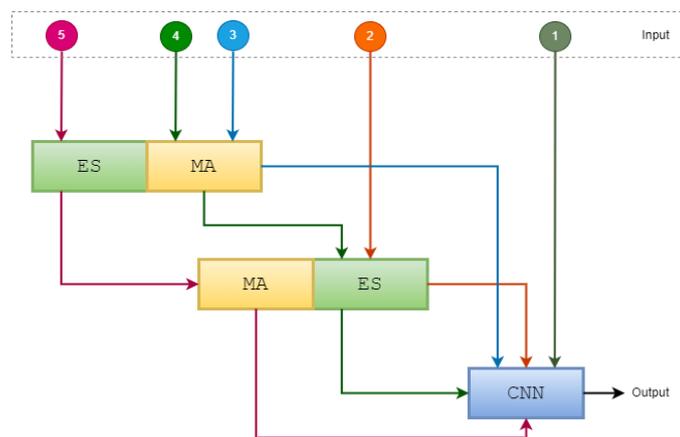


Fig. 1. Experimental Design

The use of single exponential smoothing by itself is not sufficient since it has the drawback of not being appropriate for anticipating data in seasonal and long-term periods, and the accuracy obtained is still inadequate. This makes the use of single exponential smoothing insufficient. In light of this, the motivation for the hybridization of smoothing techniques derived from single exponential smoothing with moving average (SMA) or vice versa moving average with single exponential smoothing (MAS) stems from the findings of this research.

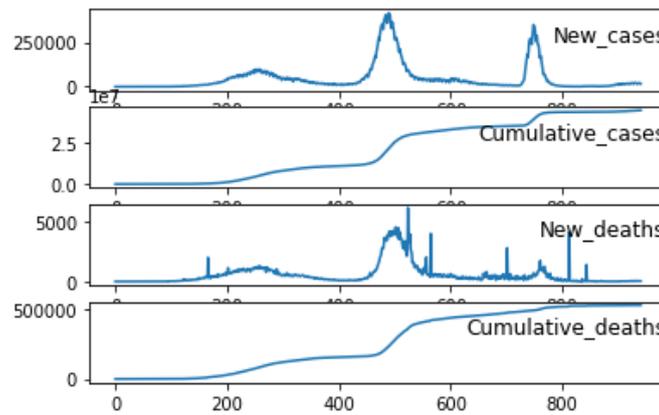
2.1. Dataset

This research used five datasets through the application of COVID-19, which had information from five countries with the highest number of instances anywhere in the globe. The dataset was obtained from the publicly available WHO website, which may be viewed and downloaded at <https://covid19.who.int/WHO-COVID-19-global-data.csv>. The WHO website is open to the general public. The information about the five countries mentioned above is included in Table 1. The time that will be looked at for this research starts on January 3, 2020, and goes through August 1, 2022.

Table 1. Big Five Countries COVID-19 Dataset

Dataset	Country	Number of Cases
1	America	90.213.060
2	India	44.036.275
3	Brazil	33.813.587
4	France	32.881.176
5	Germany	30.903.673

The new deaths attribute stores information on the number of deaths that occur each day, while the cumulative deaths attribute stores information on the total number of deaths that have occurred. The new cases attribute stores information on the number of new cases that occur each day, while the cumulative cases attribute stores the cumulative number of additional cases. The data utilized in this research is illustrated in Fig. 2, which provides a visual representation of the data.

**Fig. 2.** Visualization of Dataset

2.2. Data Normalization

Scaling a character into a particular range required by the activation function can only be accomplished through data normalization, an essential component of CNN [35]. The process of data normalization is utilized to address this issue. Since one of the primary goals of data normalization is to ensure the quality of the data before it is given to any model, its impact on the performance of any model is significant. The Min-Max normalization method was utilized in this research. Even though it is ineffective in dealing with outliers, the technique ensures that all characteristics have the same scale. The Min-Max formula is shown in (1), which produces normalized data with smaller intervals that fall inside the range 0–1 [36].

$$x' = \frac{(x - \min_x)}{(\max_x - \min_x)} \quad (1)$$

x' is the outcome of normalizing the data, x is the data that has to be normalized, \min_x is the minimum value of all the data, and \max_x is the maximum value of all the data.

2.3. Forecasting Process

In this study, the dataset used for predicting with CNN is first smoothed using single exponential smoothing and moving averages. Single exponential smoothing is employed. One may observe the equation for single exponential smoothing in (2).

$$S_t = S_{t-1} + \alpha(X_t - S_{t-1}) \quad (2)$$

The smoothed data S_t is a result of smoothing the raw data $\{X_t\}$. The smoothing factor α , is a variable that specifies the smoothing level [37]. The interval for α is between 0 and 1 ($0 \leq \alpha \leq 1$) [38]. When α is close to 1, the learning process is accelerated because the smoothing effect is diminished. In contrast, values of α closer to 0 have a greater smoothing effect and are less sensitive to changes in the

recent past [39]. Not all cases have the same value for α . Therefore, we determine the optimal dataset's properties smoothing factor value based on the dataset's properties [40]. The optimal alpha for single exponential smoothing is derived from (3). Then there is no need to manually attempt each α value from 0 to 1.

$$ptimum \alpha = \frac{(X_{max} - X_{min}) - \frac{1}{n} \sum_{i=1}^n X_t}{X_{max} - X_{min}} \tag{3}$$

So that the optimum single exponential smoothing (S_t) to improve the CNN algorithm performance used comes from the substitution of (3) to (2) results in the following (4).

$$S_t = S_{t-1} + \frac{(X_{max} - X_{min}) - \frac{1}{n} \sum_{i=1}^n X_t}{X_{max} - X_{min}} (X_t - S_{t-1}) \tag{4}$$

In the meantime, smoothing using moving average (MA) considers all of the data and uses a somewhat extended backward period. Data from the past are never left out of the computation. However, their weight in the final result is relatively minimal due to the nature of the moving average. It can illustrate ongoing trends while simultaneously eliminating fluctuations thanks to noise reduction. The data is smoothed using a moving average of either one month or thirty days when it is smoothed using MA. It is possible to visualize the MA in (5).

$$MA = \frac{(x_1 + x_2 + \dots + x_n)}{n} \tag{5}$$

MA is the outcome of smoothing the data using x , where x is the definition of each data point, and n is the number of periods.

The CNN algorithm is the primary focus of this research. CNN employs the fundamental Neural Network (NN) algorithm with additional layers. Because of its effectiveness, CNN has garnered much attention in computer vision and image processing. CNN uses a convolution layer that can process the spatial information in images, while fully connected layers are equipped with a memory that allows them to store information from time-series data. The input given to the model, an image matrix for computer vision problems and a 1D array for time series forecasting, is the only thing that differentiates computer vision problems from time series problems. The observation sequence can treat the raw input data as a one-dimensional array, which the CNN model can then read and filter. Therefore, the use of this theory in the time-series analysis is possible. CNN architecture as shown in Fig. 3

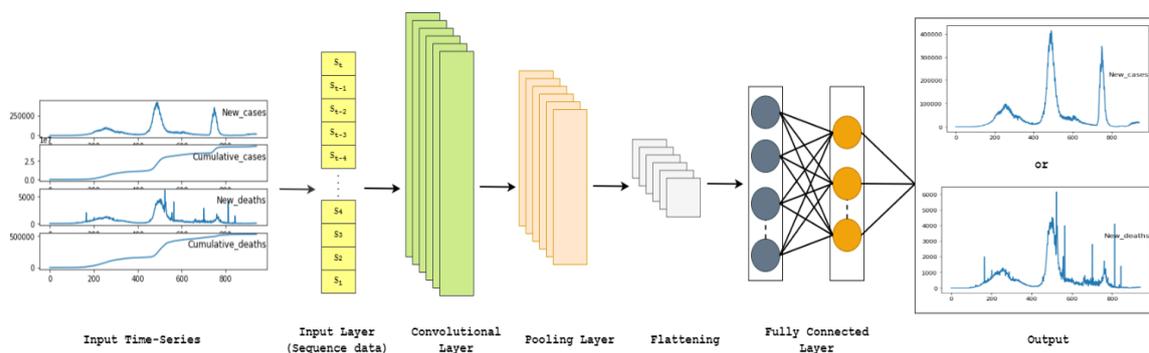


Fig. 3. CNN architecture

CNN's architecture comprises the following layers: an input layer, a convolutional layer, a pooling layer, a flattening, and a fully connected layer, as well as an output layer. Convolutional and pooling layers are designed to filter input data and extract valuable information for a fully connected network layer. Convolutional layers use raw input data and kernels to create new feature values. This technique was designed to extract image features from structured matrices. The convolution kernel (filter) is a narrow window containing coefficient values in matrix form. All these procedures result in a convolved

matrix representing a feature value specified by the filter coefficients and dimension size. By applying alternative convolution kernels to the input data, additional convolved features can be formed, usually more helpful than the original beginning features, boosting the model's performance.

A nonlinear activation function follows convolutional layers. Two typical activation functions are the sigmoid function and the rectified linear unit (ReLU). Both can be stated using (6) and (7) [41]. A pooling layer subsamples convolved features to create a lower-dimensional matrix. As with the convolutional layer, the pooling layer uses a small sliding window to take the values of each patch of convolved features and output one new value. Maximum and average pooling calculate each patch's maximum and average values. The pooling layer creates additional matrices that summarize the convolutional layer's features. Small input changes will not affect pooled output values, making the system more robust.

$$\sigma(z) = \frac{1}{1+e^{-z}} \quad (6)$$

$$f(x) = \max(0, x) \quad (7)$$

A list of CNN parameters can be adjusted in several different ways, depending on the application. Research [40] provides the basis for establishing the CNN parameters used in this investigation. In this work, we modified the parameter settings in the fully connected layer by optimizing the hyperparameter tuning using particle swarm optimization (PSO) [42], was done so that everything would not be precisely the same. The main reason is that the fully connected layer of the CNN reflects a more comprehensive set of features than the convolution layer. Each neuron in a fully connected layer is connected to all of the neurons in the layer below it [43]. A list of CNN forecast component parameters can be seen in Table 2.

Table 2. The list of CNN forecast component parameters

Layer	Parameter	Value
Convolutional	Type of convolutional	Conv1D
	The number of convolutional	1
	The number of filters	32
	The filter size	1
Pooling	The activation function	ReLU
	Type of pooling	MaXPooling1D
	The number of pooling layer	1
	Size of the pooling window	1
Flattent	Dropout	0.2
	The number of flatten layer	1
*Fully connected	The number of hidden layers	2
	The number of units or neuron	64
	The Activation function output	ReLU
	Loss function	MSE
	Type of optimizer	Adam
	The number of epochs	100
	The batch size	64

2.4. Key Performance Indicator

All experiments in this research were evaluated using key performance indicators, mean absolute percentage error (MAPE), and mean square error (MSE). In order to display errors in a manner that indicates accuracy, the MAPE metric is applied [27]. The MSE is a metric that can be used to detect outliers in a prediction system that has been created [44]. The MAPE and MSE values should be lower for a more accurate results prediction. The equation of MAPE and MSE can be seen in (8) and (9) [45].

$$MAPE = \sum_{t=1}^n \frac{|(A_t - F_t)|}{n \cdot A_t} \times 100 \quad (8)$$

$$MSE = \sum_{t=1}^n \frac{(A_t - F_t)^2}{n} \quad (9)$$

3. Results and Discussion

All of the data had been smoothed down by the experimental design of this research. At the same time, CNN was applied to the data. The outcomes of the forecasting evaluation were obtained in the form of MAPE and MSE, which can be viewed in [Table 3](#) and [Table 4](#).

Table 3. MAPE of New Cases

Method	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5
CNN	10.90335	9.98161	10.56073	9.78018	10.87463
S-CNN	10.48021	9.73711	9.07266	9.54466	10.77463
MA-CNN	10.46680	9.41301	9.05266	9.48466	10.71020
SMA-CNN	10.38977	9.08656	9.03067	9.08466	10.28978
MAS-CNN	10.53081	9.52832	9.05067	9.22466	10.74021

The MAPE for the new case forecasts is presented in [Table 3](#). According to the table, we can conclude that the performance of CNN can be enhanced by utilizing exponential smoothing (S), moving average (MA), and its combination. Because of the treatment, there is a decline in the value of the MAPE. When applied to Dataset 3, SMA-CNN achieves the most outstanding performance of 9.03067. While using CNN, the worst possible acceptable result is 10.90335 for Dataset 1. As a result, SMA-CNN is the most effective algorithm for forecasting across all of the datasets included in this research.

Table 4. MSE of New Cases

Method	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5
CNN	0.00283	0.00337	0.00492	0.00431	0.00593
S-CNN	0.00252	0.00336	0.00432	0.00381	0.00445
MA-CNN	0.00216	0.00290	0.00402	0.00193	0.00443
SMA-CNN	0.00199	0.00166	0.00362	0.00185	0.00299
MAS-CNN	0.00257	0.00298	0.00382	0.00225	0.00312

There is a connection between sensitivity and the application of MSE in performance testing. According to [Table 4](#), the MSE is reduced when the utilized algorithm is more complicated. In other words, SMA-CNN has a higher sensitivity than any other CNN-based algorithm that was tested for this study. Because the value of MSE is not very high, it can be deduced that the algorithm under consideration can recognize anomalies within the datasets used for forecasting. The results shown in [Table 5](#).

Table 5. MAPE of New Deaths

Method	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5
CNN	10.59285	9.84640	9.63077	9.68018	10.97463
S-CNN	10.53449	9.31762	9.12146	9.67486	10.87463
MA-CNN	10.50479	9.20877	9.03994	9.21712	10.77463
SMA-CNN	10.33826	9.09776	9.02606	9.09712	10.38479
MAS-CNN	10.57165	9.24941	9.03624	9.13712	10.83463

[Table 5](#) are comparable to those in [Table 3](#). The MAPE that is the lowest is 9.02606, while the MAPE that is the greatest is 10.97463 (Dataset 5 using CNN). In every dataset, the performance of SMA-CNN was superior to that of CNN, S-CNN, MA-CNN, and MAS-CNN. The value reflects the

consistency of the accuracy of the forecast, although, under the same situations (Datasets 1, 2, and 5), MA-CNN performs somewhat better than MAS-CNN does. MSE of New Deats as show in Table 6.

Table 6. MSE of New Deats

Method	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5
CNN	0.00602	0.00307	0.00832	0.00257	0.00373
S-CNN	0.00531	0.00302	0.00553	0.00194	0.00324
MA-CNN	0.00518	0.00216	0.00323	0.01390	0.00285
SMA-CNN	0.00456	0.00197	0.00263	0.00174	0.00209
MAS-CNN	0.00536	0.00243	0.00303	0.00214	0.00305

Table 6 has a similar pattern to Table 4. According to the data presented in the table, SMA-CNN has the smallest MSE, with 0.00174 being the best possible result (Dataset 4). This score, which is very near zero, indicates that all CNN variations can make accurate predictions. The MSE for SMA-CNN is the lowest of all the datasets.

S-CNN performs more accurately than CNN in most elements of the comparison, resulting from the application of optimal alpha, which was carried out [40]. The optimal value of alpha compels the process of smoothing to arrive at its optimal state, which results in a result that is both quick and accurate. In addition, the PSO hyperparameter tuning can produce an ideal model by applying the algorithm, thereby reducing errors in the dataset [42]. This smoothing process focuses more on values with the moving average timing in a given period, which can make the data more stable. MA-CNN may also improve results because of this smoothing process's attention to these values. It has been demonstrated that the smoothing technique known as smoothing moving average (SMA) has a performance that, when combined with exponential smoothing and moving average, can make time-series data prone to high volatility more stable.

On the other hand, MAS-CNN has been shown to have worse performance than MA-CNN in several tests. The series has no choice but to take on a linear form due to the exponential smoothing process after the moving average. Because of this condition, the absolute and mean square errors may increase. As a result of the MAPE and MSE, the SMA-CNN is considered the most practical combination of the moving average and exponential smoothing. We can use other exponential smoothings for further research, such as double exponential smoothing and triple exponential smoothing.

4. Conclusion

This research aims to improve the efficiency of CNN, an algorithm frequently utilized for image processing, by applying a smoothing strategy to its time-series analysis. Based on the investigation findings, one may conclude that the SMA-CNN model with the optimal smoothing factor performs significantly better than the other CNN-based forecasting smoothing technique techniques. The SMA-CNN model used in this investigation yields the highest-quality assessment results. The usage of moving averages in combination with single exponential smoothing is continued as a data preparation strategy since it considerably improves the effectiveness of the forecasting algorithm. Although the results of this research have addressed the study's objectives, there are still some limitations. The implementation of smoothing techniques that are optimized in CNN methods is the primary subject of this study. As a result, research will be conducted shortly to determine how applying this strategy to more sophisticated deep learning algorithms (such as LSTM, DBN, and RBF) would affect the results. The next item we will concentrate on is conducting a more in-depth study of the various smoothing methods that may be implemented using double or triple-exponential smoothing. In the future study that is conducted, both beta optimal and gamma optimum will be taken into consideration.

Acknowledgment

This work was supported by Laboratory of Artificial Intelligence, Universitas Mulawarman, Indonesia and Laboratory of Machine Learning, Universiti Malaysia Sabah, Malaysia.

Declarations

Author contribution. All authors contributed equally as the main contributor of this paper. All authors read and approved the final paper.

Funding statement. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest. The authors declare no conflict of interest.

Additional information. No additional information is available for this paper.

References

- [1] S. Zhao *et al.*, "Imitation dynamics in the mitigation of the novel coronavirus disease (COVID-19) outbreak in Wuhan, China from 2019 to 2020," *Ann. Transl. Med.*, vol. 8, no. 7, pp. 448–448, Apr. 2020, doi: [10.21037/atm.2020.03.168](https://doi.org/10.21037/atm.2020.03.168).
- [2] S. Chahar and P. K. Roy, "COVID-19: A Comprehensive Review of Learning Models," *Arch. Comput. Methods Eng.*, vol. 29, no. 3, pp. 1915–1940, May 2022, doi: [10.1007/s11831-021-09641-3](https://doi.org/10.1007/s11831-021-09641-3).
- [3] S. S. Diwan, S. Ravichandran, R. Govindarajan, and R. Narasimha, "Understanding Transmission Dynamics of COVID-19-Type Infections by Direct Numerical Simulations of Cough/Sneeze Flows," *Trans. Indian Natl. Acad. Eng.*, vol. 5, no. 2, pp. 255–261, Jun. 2020, doi: [10.1007/s41403-020-00106-w](https://doi.org/10.1007/s41403-020-00106-w).
- [4] L. A. Abraham, T. C. Brown, and S. A. Thomas, "How COVID-19's Disruption of the U.S. Correctional System Provides an Opportunity for Decarceration," *Am. J. Crim. Justice*, vol. 45, no. 4, pp. 780–792, Aug. 2020, doi: [10.1007/s12103-020-09537-1](https://doi.org/10.1007/s12103-020-09537-1).
- [5] M. Mirbabaie, S. Stieglitz, and N. R. J. Frick, "Artificial intelligence in disease diagnostics: A critical review and classification on the current state of research guiding future direction," *Health Technol. (Berl.)*, vol. 11, no. 4, pp. 693–731, Jul. 2021, doi: [10.1007/s12553-021-00555-5](https://doi.org/10.1007/s12553-021-00555-5).
- [6] S. Kaur *et al.*, "Medical Diagnostic Systems Using Artificial Intelligence (AI) Algorithms: Principles and Perspectives," *IEEE Access*, vol. 8, pp. 228049–228069, 2020, doi: [10.1109/ACCESS.2020.3042273](https://doi.org/10.1109/ACCESS.2020.3042273).
- [7] R. Liu, B. Yang, E. Zio, and X. Chen, "Artificial intelligence for fault diagnosis of rotating machinery: A review," *Mech. Syst. Signal Process.*, vol. 108, pp. 33–47, Aug. 2018, doi: [10.1016/j.ymssp.2018.02.016](https://doi.org/10.1016/j.ymssp.2018.02.016).
- [8] M. Jamshidi *et al.*, "Artificial Intelligence and COVID-19: Deep Learning Approaches for Diagnosis and Treatment," *IEEE Access*, vol. 8, pp. 109581–109595, 2020, doi: [10.1109/ACCESS.2020.3001973](https://doi.org/10.1109/ACCESS.2020.3001973).
- [9] C. Cuello-Garcia, G. Pérez-Gaxiola, and L. van Amelsvoort, "Social media can have an impact on how we manage and investigate the COVID-19 pandemic," *J. Clin. Epidemiol.*, vol. 127, pp. 198–201, Nov. 2020, doi: [10.1016/j.jclinepi.2020.06.028](https://doi.org/10.1016/j.jclinepi.2020.06.028).
- [10] S. Zeadally, E. Adi, Z. Baig, and I. A. Khan, "Harnessing Artificial Intelligence Capabilities to Improve Cybersecurity," *IEEE Access*, vol. 8, pp. 23817–23837, 2020, doi: [10.1109/ACCESS.2020.2968045](https://doi.org/10.1109/ACCESS.2020.2968045).
- [11] R. Vaishya, M. Javaid, I. H. Khan, and A. Haleem, "Artificial Intelligence (AI) applications for COVID-19 pandemic," *Diabetes Metab. Syndr. Clin. Res. Rev.*, vol. 14, no. 4, pp. 337–339, Jul. 2020, doi: [10.1016/j.dsx.2020.04.012](https://doi.org/10.1016/j.dsx.2020.04.012).
- [12] A. Holzinger, P. Kieseberg, E. Weippl, and A. M. Tjoa, "Current Advances, Trends and Challenges of Machine Learning and Knowledge Extraction: From Machine Learning to Explainable AI," in *Lecture Notes in Computer Science*, 2018, pp. 1–8, doi: [10.1007/978-3-319-99740-7_1](https://doi.org/10.1007/978-3-319-99740-7_1).
- [13] A. Ignatov *et al.*, "AI Benchmark: All About Deep Learning on Smartphones in 2019," in *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, Oct. 2019, pp. 3617–3635, doi: [10.1109/ICCVW.2019.00447](https://doi.org/10.1109/ICCVW.2019.00447).
- [14] Y.-J. Cao *et al.*, "Recent Advances of Generative Adversarial Networks in Computer Vision," *IEEE Access*, vol. 7, pp. 14985–15006, 2019, doi: [10.1109/ACCESS.2018.2886814](https://doi.org/10.1109/ACCESS.2018.2886814).
- [15] Y. Sun, B. Pan, and Y. Fu, "Lightweight Deep Neural Network for Real-Time Instrument Semantic Segmentation in Robot Assisted Minimally Invasive Surgery," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 3870–3877, Apr. 2021, doi: [10.1109/LRA.2021.3066956](https://doi.org/10.1109/LRA.2021.3066956).
- [16] X. Zhang *et al.*, "Deep Learning Based Analysis of Breast Cancer Using Advanced Ensemble Classifier and

- Linear Discriminant Analysis,” *IEEE Access*, vol. 8, pp. 120208–120217, 2020, doi: [10.1109/ACCESS.2020.3005228](https://doi.org/10.1109/ACCESS.2020.3005228).
- [17] J. Sun, P. Veeffkind, P. van Velthoven, and P. F. Levelt, “Aerosol Absorption Over Land Derived From the Ultra-Violet Aerosol Index by Deep Learning,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 14, pp. 9692–9710, 2021, doi: [10.1109/JSTARS.2021.3108669](https://doi.org/10.1109/JSTARS.2021.3108669).
- [18] X.-H. Le, D.-H. Nguyen, S. Jung, M. Yeon, and G. Lee, “Comparison of Deep Learning Techniques for River Streamflow Forecasting,” *IEEE Access*, vol. 9, pp. 71805–71820, 2021, doi: [10.1109/ACCESS.2021.3077703](https://doi.org/10.1109/ACCESS.2021.3077703).
- [19] A. Mahmoud and A. Mohammed, “A Survey on Deep Learning for Time-Series Forecasting,” in *Machine Learning and Big Data Analytics Paradigms: Analysis, Applications and Challenges*, A. Hassanien Aboul Ella and Darwish, Ed., Cham: Springer International Publishing, 2021, pp. 365–392, doi: [10.1007/978-3-030-59338-4_19](https://doi.org/10.1007/978-3-030-59338-4_19).
- [20] W. Zhong, N. Yu, and C. Ai, “Applying big data based deep learning system to intrusion detection,” *Big Data Min. Anal.*, vol. 3, no. 3, pp. 181–195, Sep. 2020, doi: [10.26599/BDMA.2020.9020003](https://doi.org/10.26599/BDMA.2020.9020003).
- [21] S. Bahri, M. Kdayem, and N. Zoghlami, “Long Short-Term Memory based RNN for COVID-19 disease prediction,” in *2021 22nd IEEE International Conference on Industrial Technology (ICIT)*, Mar. 2021, pp. 901–906, doi: [10.1109/ICIT46573.2021.9453534](https://doi.org/10.1109/ICIT46573.2021.9453534).
- [22] P. Wang, X. Zheng, G. Ai, D. Liu, and B. Zhu, “Time series prediction for the epidemic trends of COVID-19 using the improved LSTM deep learning method: Case studies in Russia, Peru and Iran,” *Chaos, Solitons & Fractals*, vol. 140, p. 110214, Nov. 2020, doi: [10.1016/j.chaos.2020.110214](https://doi.org/10.1016/j.chaos.2020.110214).
- [23] Y. Fei, “Analysis the use of machine learning algorithm-based methods in predicting COVID-19 infection,” in *Proceedings of the 2nd International Symposium on Artificial Intelligence for Medicine Sciences*, Oct. 2021, pp. 88–94, doi: [10.1145/3500931.3500948](https://doi.org/10.1145/3500931.3500948).
- [24] L. Alzubaidi *et al.*, “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions,” *J. Big Data*, vol. 8, no. 1, p. 53, Dec. 2021, doi: [10.1186/s40537-021-00444-8](https://doi.org/10.1186/s40537-021-00444-8).
- [25] L. Ren, J. Dong, X. Wang, Z. Meng, L. Zhao, and M. J. Deen, “A Data-Driven Auto-CNN-LSTM Prediction Model for Lithium-Ion Battery Remaining Useful Life,” *IEEE Trans. Ind. Informatics*, vol. 17, no. 5, pp. 3478–3487, May 2021, doi: [10.1109/TII.2020.3008223](https://doi.org/10.1109/TII.2020.3008223).
- [26] T. Cheng, F. Harrou, F. Kadri, Y. Sun, and T. Leiknes, “Forecasting of Wastewater Treatment Plant Key Features Using Deep Learning-Based Models: A Case Study,” *IEEE Access*, vol. 8, pp. 184475–184485, 2020, doi: [10.1109/ACCESS.2020.3030820](https://doi.org/10.1109/ACCESS.2020.3030820).
- [27] S. Karasu, A. Altan, S. Bekiros, and W. Ahmad, “A new forecasting model with wrapper-based feature selection approach using multi-objective optimization technique for chaotic crude oil time series,” *Energy*, vol. 212, p. 118750, Dec. 2020, doi: [10.1016/j.energy.2020.118750](https://doi.org/10.1016/j.energy.2020.118750).
- [28] M. B. A. Rabbani *et al.*, “A Comparison Between Seasonal Autoregressive Integrated Moving Average (SARIMA) and Exponential Smoothing (ES) Based on Time Series Model for Forecasting Road Accidents,” *Arab. J. Sci. Eng.*, vol. 46, no. 11, pp. 11113–11138, Nov. 2021, doi: [10.1007/s13369-021-05650-3](https://doi.org/10.1007/s13369-021-05650-3).
- [29] S. Smyl, “A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting,” *Int. J. Forecast.*, vol. 36, no. 1, pp. 75–85, Jan. 2020, doi: [10.1016/j.ijforecast.2019.03.017](https://doi.org/10.1016/j.ijforecast.2019.03.017).
- [30] R. K. M. Malhi *et al.*, “Applicability of Smoothing Techniques in Generation of Phenological Metrics of *Tectona grandis* L. Using NDVI Time Series Data,” *Remote Sens.*, vol. 13, no. 17, p. 3343, Aug. 2021, doi: [10.3390/rs13173343](https://doi.org/10.3390/rs13173343).
- [31] O. Mbaye, S. Konare, M. L. Ba, and A. Diop, “Short-Term Holt Smoothing Prediction Model of Daily COVID-19 Reported Cumulative Cases,” in *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, 2021, pp. 15–27, doi: [10.1007/978-3-030-90556-9_2](https://doi.org/10.1007/978-3-030-90556-9_2).
- [32] X. Yang and Y. Ni, “Least-squares estimation for uncertain moving average model,” *Commun. Stat. - Theory Methods*, vol. 50, no. 17, pp. 4134–4143, Sep. 2021, doi: [10.1080/03610926.2020.1713373](https://doi.org/10.1080/03610926.2020.1713373).
- [33] J. F. Rendon-Sanchez and L. M. de Menezes, “Structural combination of seasonal exponential smoothing forecasts applied to load forecasting,” *Eur. J. Oper. Res.*, vol. 275, no. 3, pp. 916–924, Jun. 2019, doi:

[10.1016/j.ejor.2018.12.013](https://doi.org/10.1016/j.ejor.2018.12.013).

- [34] M. Li, X. Liu, and F. Ding, "The filtering-based maximum likelihood iterative estimation algorithms for a special class of nonlinear systems with autoregressive moving average noise using the hierarchical identification principle," *Int. J. Adapt. Control Signal Process.*, vol. 33, no. 7, pp. 1189–1211, Jul. 2019, doi: [10.1002/acs.3029](https://doi.org/10.1002/acs.3029).
- [35] M. Yedroudj, F. Comby, and M. Chaumont, "Yedroudj-Net: An Efficient CNN for Spatial Steganalysis," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2018, pp. 2092–2096, doi: [10.1109/ICASSP.2018.8461438](https://doi.org/10.1109/ICASSP.2018.8461438).
- [36] A. R. F. Dewandra, A. P. Wibawa, U. Pujiyanto, A. B. P. Utama, and A. Nafalski, "Journal Unique Visitors Forecasting Based on Multivariate Attributes Using CNN," *Int. J. Artif. Intell. Res.*, vol. 6, no. 1, 2022, doi: [10.29099/ijair.v6i1.274](https://doi.org/10.29099/ijair.v6i1.274).
- [37] C. Ning and F. You, "Data-driven decision making under uncertainty integrating robust optimization with principal component analysis and kernel smoothing methods," *Comput. Chem. Eng.*, vol. 112, pp. 190–210, Apr. 2018, doi: [10.1016/j.compchemeng.2018.02.007](https://doi.org/10.1016/j.compchemeng.2018.02.007).
- [38] J. Mi, L. Fan, X. Duan, and Y. Qiu, "Short-Term Power Load Forecasting Method Based on Improved Exponential Smoothing Grey Model," *Math. Probl. Eng.*, vol. 2018, pp. 1–11, 2018, doi: [10.1155/2018/3894723](https://doi.org/10.1155/2018/3894723).
- [39] C. Pan, Y. Chen, L. Wang, and Z. He, "Lithium-ion Battery Remaining Useful Life Prediction Based on Exponential Smoothing and Particle Filter," *Int. J. Electrochem. Sci.*, vol. 14, pp. 9537–9551, Oct. 2019, doi: [10.20964/2019.10.15](https://doi.org/10.20964/2019.10.15).
- [40] A. P. Wibawa, A. B. P. Utama, H. Elmunsyah, U. Pujiyanto, F. A. Dwiyanto, and L. Hernandez, "Time-series analysis with smoothed Convolutional Neural Network," *J. Big Data*, vol. 9, no. 1, p. 44, Dec. 2022, doi: [10.1186/s40537-022-00599-y](https://doi.org/10.1186/s40537-022-00599-y).
- [41] Z. Shen, X. Fan, L. Zhang, and H. Yu, "Wind speed prediction of unmanned sailboat based on CNN and LSTM hybrid neural network," *Ocean Eng.*, vol. 254, p. 111352, Jun. 2022, doi: [10.1016/j.oceaneng.2022.111352](https://doi.org/10.1016/j.oceaneng.2022.111352).
- [42] A. Pranolo, Y. Mao, A. P. Wibawa, A. B. P. Utama, and F. A. Dwiyanto, "Robust LSTM With Tuned-PSO and Bifold-Attention Mechanism for Analyzing Multivariate Time-Series," *IEEE Access*, vol. 10, pp. 78423–78434, 2022, doi: [10.1109/ACCESS.2022.3193643](https://doi.org/10.1109/ACCESS.2022.3193643).
- [43] S. Liu, H. Ji, and M. C. Wang, "Nonpooling Convolutional Neural Network Forecasting for Seasonal Time Series With Trends," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 31, no. 8, pp. 2879–2888, Aug. 2020, doi: [10.1109/TNNLS.2019.2934110](https://doi.org/10.1109/TNNLS.2019.2934110).
- [44] D. J. Kovacs *et al.*, "Membrane fouling prediction and uncertainty analysis using machine learning: A wastewater treatment plant case study," *J. Memb. Sci.*, vol. 660, p. 120817, Oct. 2022, doi: [10.1016/j.memsci.2022.120817](https://doi.org/10.1016/j.memsci.2022.120817).
- [45] S. Khullar and N. Singh, "Water quality assessment of a river using deep learning Bi-LSTM methodology: forecasting and validation," *Environ. Sci. Pollut. Res.*, vol. 29, no. 9, pp. 12875–12889, Feb. 2022, doi: [10.1007/s11356-021-13875-w](https://doi.org/10.1007/s11356-021-13875-w).