An accurate traffic flow prediction using long-short term memory and gated recurrent unit networks

Mohamed S. Sawah[1], Shereen Aly Taie[1], Mohamed Hasan Ibrahim[2], Shereen A. Hussein[1]
[1]Department of Computer Science, Faculty of Computers and Artificial Intelligence, Fayoum University, Fayoum, Egypt
[2]Department of Information Systems, Faculty of Computers and Artificial Intelligence, Fayoum University, Fayoum, Egypt

Article Info

Article history:

Received Dec 14, 2022
Revised Dec 25, 2022
Accepted Jan 10, 2023

Keywords:

Deep learning
Survellience camera
Traffic congestion
Traffic flow
Traffic management

ABSTRACT

Congestion on roadways is an issue in many cities, especially at peak times, which causes air and noise pollution and cause pressure on citizens. So, the implementation of intelligent transportation systems (ITSs) is a very important part of smart cities. As a result, the importance of making accurate short-term predictions of traffic flow has significantly increased in recent years. However, the current methods for predicting short-term traffic flow are incapable of effectively capturing the complex non-linearity of traffic flow that affects prediction accuracy. To overcome this problem, this study introduces two novel models. The first model uses two long-short term memory (LSTM) units that can extract the traffic flow temporal features followed by four dense layers to perform the traffic flow prediction. The second model uses two gated recurrent unit (GRU) units that can extract the traffic flow temporal features followed by three dense layers to perform the traffic flow prediction. The two proposed models give promising results on performance measurement system (PEMS), traffic and congestions (TRANCOS) dataset that is firstly used as metadata. So, the two models can do this in specific cases and are able to suddenly capture trend changes.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Mohamed S. Sawah
Department of Computer Science, Faculty of Computers and Artificial Intelligence, Fayoum University
Fayoum, Egypt
Email: me1900@fayoum.edu.eg

1. INTRODUCTION

The importance of accurate forecasting of traffic in the short term lies in facilitating the daily lives of citizens, as it provides them with the best paths to reduce time and effort during trips, in addition to helping government agencies develop appropriate plans to prevent traffic congestion [1]. So, the efficient prediction of traffic flow data has become a crucial part of intelligent transportation systems (ITSs). Although, short-term traffic flow forecasting is extremely difficult because of the non-linearity of traffic data. With the proliferation of ITSs in recent years, the ability to precisely and effectively estimate traffic flow has drawn considerable interest [2]. The prediction of Traffic flow is the process of predicting the traffic flow distribution by using the historical data of traffic flow. It consists of both short-term and long-term traffic flow forecasts. The range of short-term traffic forecasts is five minutes to one hour [3]. Real-time and reliable data on short-term traffic changes can reduce traffic strain, prevent accidents, and alleviate traffic congestion to some extent. So, short-term traffic forecasting has become an important research area in the traffic flow prediction field [4].

The growth of data collecting and sensor technologies of traffic flow data provides more robust data and modeling. The features of traffic flow data are non-linearity, periodicity, randomness, and volatility [5]. So, the prediction using traditional methods is very difficult such as autoregressive integrated moving average (ARIMA), support vector regression (SVR), and multi-variable linear regression. These methods do not account for the full range of traffic flow features and so do not achieve accurate traffic forecasting [6]. Continuous advancements in machine learning and deep learning theory create new challenges for ITS development. Due to their robust non-linear fitting capabilities, deep learning networks are commonly employed for short-term traffic flow predictions [7].

This study introduces two novel models. The first model uses two long-short term memory (LSTM) units that can extract the traffic flow temporal features followed by four dense layers to perform the traffic flow prediction. The second model uses two gated recurrent unit (GRU) units that can extract the traffic flow temporal features followed by three dense layers to perform the traffic flow prediction. The main advantage of these models its ability to effectively capture the complex non-linearity of traffic flow, which will improve the traffic flow prediction accuracy. The structure of the paper is coming as section 2 presents the related works for traffic flow prediction, section 3 presents the proposed method, the experimental work and results are explained in section 4, and the paper is concluded in section 5.

2. RELATED WORK

The increasing volume of traffic has impacted the viability of urban growth. A lot of strategies have been investigated for forecasting traffic flow with great accuracy, taking into consideration a variety of factors and characteristics. There are typically three main types of traffic flow prediction techniques: parametric techniques, non-parametric techniques, and deep learning techniques.

2.1. Parametric techniques

The most commonly employed parametric model is (ARIMA) model, a time-series technique. It is standard for estimating the flow of traffic over the next few short periods of time, and it is also used to predict the flow of traffic on expressways and urban areas [8]. Numerous variations of the ARIMA model were subsequently presented to improve the performance of the prediction. For instance, a KARIMA model combining ARIMA with the Kohonen network for prediction addresses the issue of non-linearity of the data, as the ARIMA model cannot deal with this kind of data [9]. Combining ARIMA with explanatory variables, an ARIMAX model was suggested to increase forecasting performance [10]. Ghosh *et al*. [11] developed the Bayesian Seasonal ARIMA model that utilizes the Bayesian method to increase forecasting performance. Duan *et al.* [12] developed the spatio-temporal ARIMA (STARIMA) model to extract the spatio-temporal features in the data of traffic flow to achieve more precise prediction performance. Parametric models cannot describe traffic flow accurately using analysis formulas due to their non-linear stochastic nature. Therefore, there is great interest in non-parametric models in traffic flow forecasting.

2.2. Non-parametric techniques

The non-parametric models include some approaches, such as support vector regression, k-nearest neighbor (k-NN) methods, and artificial neural networks (ANNs) [13], [14]. For instance, the k-NN is used to estimate short-term traffic flow, although its performance is lower than time-series linear methods [15]. Chang *et al.* [16] proposed a model by utilizing an abundance of past data to enhance forecasting performance. Jeong *et al*. [17] suggested weighted online learning (SVR) to increase forecasting performance. In addition, numerous ANN-based models have been presented in [18]–[22] for traffic flow forecasting. Moreover, early neural network-based works typically employed shallow networks, which were incapable of capturing the uncertainty and complicated non-linearity of the traffic flow [2]. So, the non-parametric models are insufficient for achieving reliable prediction performance.

2.3. Hybrid techniques

A hybrid technique is the integration of two or more models to increase performance. Hybrid techniques are used to overcome the dynamic change in traffic flow, which results in uncertainty in traffic flow. On this basis, hybrid models have been proposed to adapt to these changes to predict the behavior of vehicles in the short term. For instance, an aggregation model is employed in [23] to achieve more precise prediction performance. For traffic flow forecasting, a method integrating the ARIMA model with cumulative sum algorithms was introduced in [24]. Dimitriou *et al.* [25] proposed an adaptive hybrid fuzzy rule-based method for traffic flow forecasting.

2.4. Deep learning techniques

Deep learning can increase the prediction performance of traffic flow, saving time and costs [26]. Due to the stochastic nature and the non-linearity of urban mobility data, deep learning is utilized in several studies to identify patterns and develop suitable forecasting models to predict urban mobility data [27], [28]. In recent years, numerous deep learning techniques have been developed for traffic flow forecasting. The first model to apply the deep learning approach is introduced in [29], where feature learning is done using the deep belief network, and the training process is done using multitask learning method.

A stacked autoencoder (SAE) model was also suggested in [30] for the prediction of traffic flow to extract the spatial and temporal characteristics of traffic data. A deep learning method employing LSTM was described in [31] to capture the temporal characteristics of traffic flow. Jia *et al.* [32] proposed a deep learning model that uses LSTM to estimate the flow of traffic in rainfall circumstances more accurately. Polson and Sokolov [33] proposed a deep learning linear model able to achieve more precise prediction performance by capturing the non-linear and the Spatio-temporal effects in the traffic data.

Du *et al.* [34] proposed a hybrid multimodal deep learning system adequately predicts complicated, non-linear urban traffic flow by incorporating modality traffic data representation properties. Xiao and Yin [35] presented a hybrid LSTM model able to deal with diverse traffic situations, and it has a smaller prediction error than other models but requires a somewhat longer running time. Wu *et al.* [36] proposed a model that uses CNN to capture the spatial features and LSTM to extract the traffic flow temporal features. Bruna *et al.* [37] proposed a model that expands CNN to a graph convolutional neural network (GCN) to learn the graph data features. Yu *et al.* [38] proposed a model for traffic flow forecasting using (GCN) able to capture the satio-temporal features in the traffic flow data. Wei *et al.* [39] proposed a model that uses the Auto Encoder to extract the upstream and downstream characteristics of traffic flow data; then, the LSTM network uses the autoencoder's acquired characteristics and previous data to achieve more precise prediction performance. Liu *et al.* [40] proposed a model to forecast the bus traffic flow. Zhang *et al.* [41] proposed a deep autoencoder model to forecast the traffic flow by extracting the temporal correlations in the traffic flow data. To overcome the issue of the complex non-linearity of traffic flow data, this study introduces two novel models. The first model uses an LSTM network and the second model uses a GRU network that can extract the temporal features of traffic flow more efficiently.

3. THE PROPOSED MODEL

3.1. Dataset

This study analyses the traffic data using two different datasets using two novel models. The first model uses an LSTM network and the second model uses a GRU network that can extract the temporal features of traffic flow more efficiently. The first dataset is performance measurement system (PEMS) collection; data are aggregated in real-time from sensors installed on the freeway in all of California's main metropolitan regions. For traffic flow forecasting, the PEMS dataset is widely used as a public benchmark. PEMS delivers real-time data from more than 39,000 sensors distributed statewide in California's freeway systems [42]. This freeway's average traffic flow is determined by averaging the traffic data obtained by several detectors. The dataset contained 7,776 training instances and 4,320 testing instances.

The second dataset is the traffic and congestions (TRANCOS) [43] dataset. It is the first dataset to count vehicles in images of traffic jams captured with real-world traffic monitoring cameras. Also, it is frequently used to assess the generalizability of vehicle counting techniques. The cameras picked monitor various motorways in the Madrid area, which are notorious for their intense traffic congestion. Each image has been annotated with a precise vehicle number and their locations for each image, where 46,796 vehicles have been annotated in total. Note that each of the collected images has traffic congestion, spanning a number of diverse scenarios and angles, with varying lighting conditions, varying degrees of crowdedness, and overlap, even in the same image. But in this study, the TRANCOS dataset is used as metadata, and this is the first time that it has been used in this form by extracting all possible features from each image in the dataset, such as (the number of vehicles, time, and date of capturing the image). The dataset has been divided into a train and test split. The train split consists of 1,031 instances, while the test split consists of 213 instances.

3.2. Model architecture

Due to the problem of non-linearity of traffic flow data. The architecture of this study introduces two novel models that deal with the non-linearity in traffic data to achieve more precise prediction performance, as shown in Figure 1. The first model is introduced based on LSTM network. The second model is introduced based on GRU network that can extract the temporal features of traffic flow more efficiently in the near future to increase transportation efficiency.

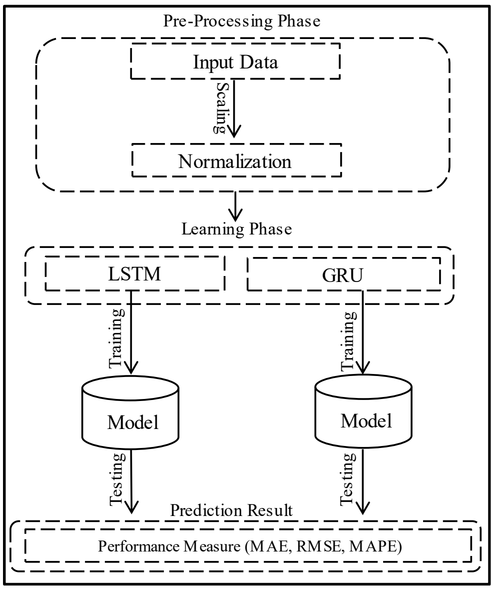

Figure 1. Model architecture diagram

3.2.1. Pre-processing phase

Data normalizing is a very crucial procedure in data preprocessing. It ensures the input data quality to deep learning models. Data normalization is necessary when features have varying value ranges. In this stage, the data normalization is used to transform features to be on a similar scale that will enhance the model's performance and training stability and reduce the time of training. There are different kinds of data normalization methods, including linear scaling, clipping, Z-score decimal scaling, and log scaling. In this study, the most popular normalization method, linear scaling to unit range normalization, is used [44].

a. Linear scalling

Linear scaling is the simplest type of scaling. It is a linear transformation technique that normalizes the data in a range, usually [0, 1] or [-1, 1]. With a lower bound min (x_i) and upper bound max (x_i) of an attribute x_i, the normalization value is given by:

$$\hat{x}_i = \frac{x_i - min\ (x_i)}{max\ (x_i) - min\ (x_i)} \tag{1}$$

3.2.2. Learning phase

The first model starts with two LSTM units that are used to extract the traffic flow temporal features. These temporal features are combined into a feature vector followed by four dense layers to perform the traffic flow prediction, as shown in Figure 2. The second model starts with two GRU units that are used to extract the traffic flow temporal features. These temporal features are combined into a feature vector followed by three dense regression layers to perform the traffic flow prediction, as shown in Figure 3. The mean squared error loss function is used as the objective function, which is discussed later. LSTM and GRU are discussed in detail in the following subsections.

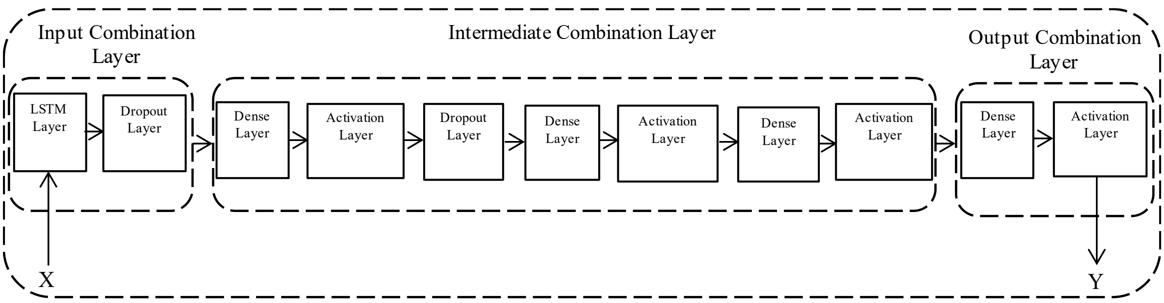

Figure 2. Model framework of LSTM

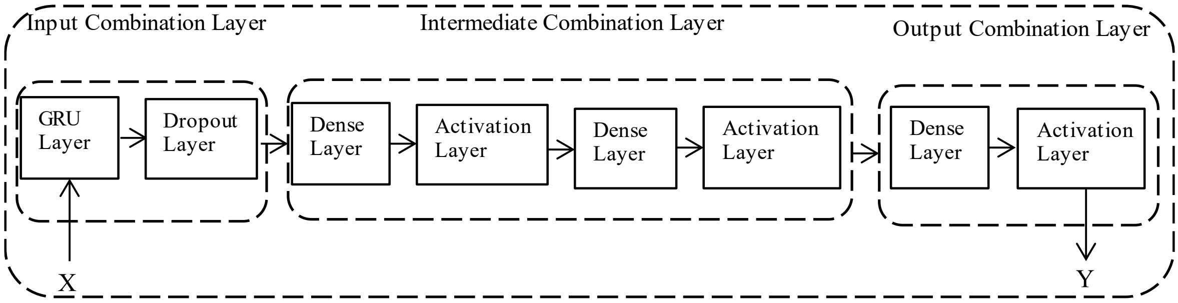

Figure 3 Model framework of GRU

a. Long-short term memory

The so-called vanishing gradient problem is present in traditional recurrent neural network (RNN) design. To overcome this disadvantage, certain RNN structures, such as LSTM, which was first introduced in [45], were built to provide memory cells the capacity to determine when to forget information, hence determining the ideal time lags for time series problems. As its enduring memory capacity, these characteristics are particularly desired for short-term traffic flow prediction in the transportation domain. The structure of LSTM NN cells is shown in Figure 4. A typical LSTM has three multiplicative units, i.e., the input gate, the forget gate, and the output gate. The input gate is used to memorize some information from the present, the forget gate is used to choose to forget some information from the past, and the output gate uses all calculated results to produce output for the LSTM NN cell [46].

In Figure 4 f_t denotes the forget gate, i_t denotes the input gate, and o_t denotes the output gate, that will together control the cellular state C_t that stores the future and historical information. The input and output information of LSTM unit is x_t and h_t respectively. The processing operation of the LSTM unit can be expressed using (2)-(7):

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{2}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

$$\tilde{C}_t = tan\,h(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{4}$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{5}$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{6}$$

$$h_t = o_t \times tan\,h(C_t) \tag{7}$$

The forget gate is represented in (2), which receives the cell state weighted sum at time t , the cell output at time $t-1$, and the activation function is the sigmoid function. The input gate is represented in (3) and has the same parameters as in (2), the memory unit value at time t is calculated in (4), and the activation function is a tanh function. The present and past memories are concatenated in (5). The output gate is represented in (6) and has the same parameters as in (2). The output of the cell is represented in (7), where the activation function is a tanh function, W is the weighted vector matrix, and b is the bias vector.

b. Gated recurrent unit

GRU is first proposed in [47], it is similar to LSTM, but it is simpler to implement and compute. The GRU cell structure is shown in Figure 5. GRU cell consists of two gates: the reset gate r and the update gate z. As in the LSTM cell, the output of the hidden state at time t is determined by the hidden state $\mathrm{t}-1$. The function of forgetting gates in LSTM is similar to the function of resetting gates in GRU. In this study, GRU NNs use the same regression part and the optimization method of the LSTM NNS, and the input time series value at time t is calculated in (8):

$$h_t = f(h_{t-1}, x_t) \tag{8}$$

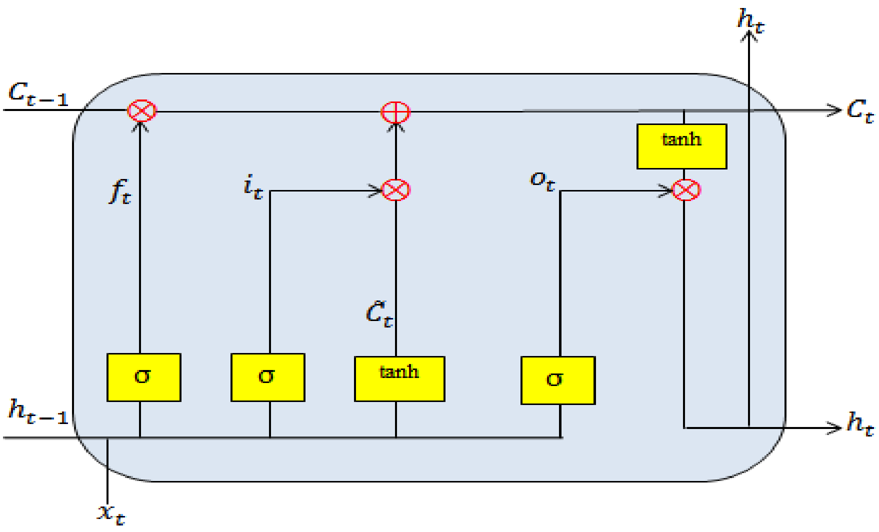

Figure 4. Structure of LSTM NN cells

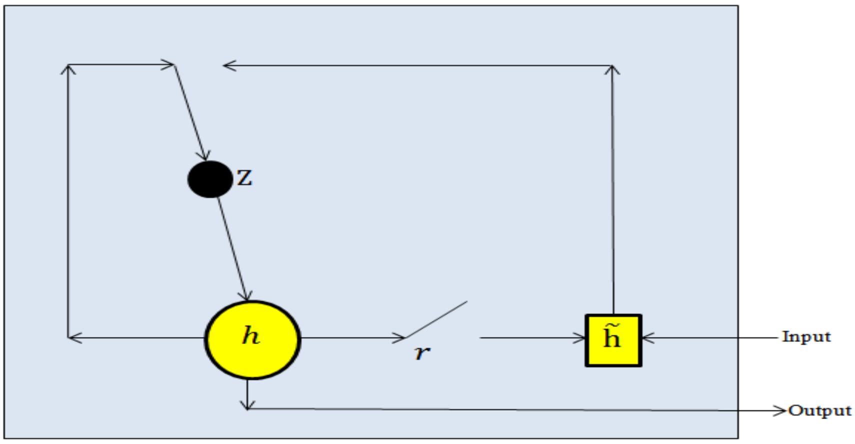

Figure 5. Structure of GRU cells

c. Training details

The two models are trained using RMSprop optimizer with a base learning rate ($\alpha = 0.0005$ based on trials and error operations). Throughout the training, a consistent learning rate was maintained, and consistent training was maintained across architectures. In addition, the mean square error loss function is used to assess the difference between the observed traffic flow and the predicted traffic flow. Here is the formula for the loss function as shown in (9):

$$L(y_t, p_t) = \frac{1}{n}\sum_{t=1}^{n} (y_t - p_t)^2 \tag{9}$$

Where the number of training examples is denoted as n and the observed traffic flow is denoted as y_t And the predicted traffic flow is denoted as p_t.

3.2.3. Prediction result

- Performance measures

The efficiency of the two models is evaluated for prediction using the following three performance measures. The first measure is the mean absolute error (MAE) which is useful for continuous variable data. The second measure is the root means square error (RMSE), the standard deviation of residuals (prediction errors). The third measure is the mean absolute percentage error (MAPE) which measures the accuracy of the prediction typically shown as a percentage. They are formulated as:

Mean absolute error:

$$MAE = \frac{1}{n}\sum_{t=1}^{n} |y_t - p_t| \tag{10}$$

Root mean square error:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} |y_t - p_t|^2} \tag{11}$$

Mean absolute percentage error:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} \frac{|y_t - p_t|}{y_t} \tag{12}$$

where $\mathrm{y_t}$ is the observed traffic flow, and $\mathrm{p_t}$ is the predicted traffic flow.

4. EXPERIMENTAL WORK AND RESULTS

The experiments are performed on two traffic flow datasets: PEMS [42] and TRANCOS [43] dataset. The two models are trained using a rmsprop optimizer with a base learning rate ($\alpha = 0.0005$ based on trials and error operations). Throughout the training, a consistent learning rate was maintained, and consistent training was maintained across architectures. In Figure 2 and Figure 3, the input combination layer and the intermediate combination layer, and the output combination layer has the same activation function, and the activation function is RELU activation function. The three combination layers have the same dropout rate and are set to 0.2. The experimental results indicate that the two proposed models give a promising result, indicating that they can perform well in specific cases and are able to suddenly capture trend changes in the traffic data flow.

According to Figure 6, GRU introduces a significant performance on the PEMS dataset that will give an accurate prediction of the traffic flow at most of the predicted times. According to Figure 7, LSTM introduces a significant performance on the TRANCOS dataset that will give an accurate prediction of the traffic flow at most of the predicted times. As shown in Table 1, the two models give a promising result on the PEMS dataset in terms of MAE, MAPE, and RMSE. As shown in Table 2, the two models give a promising result on the TRANCOS dataset that is firstly used as metadata in terms of MAE, MAPE, and RMSE.

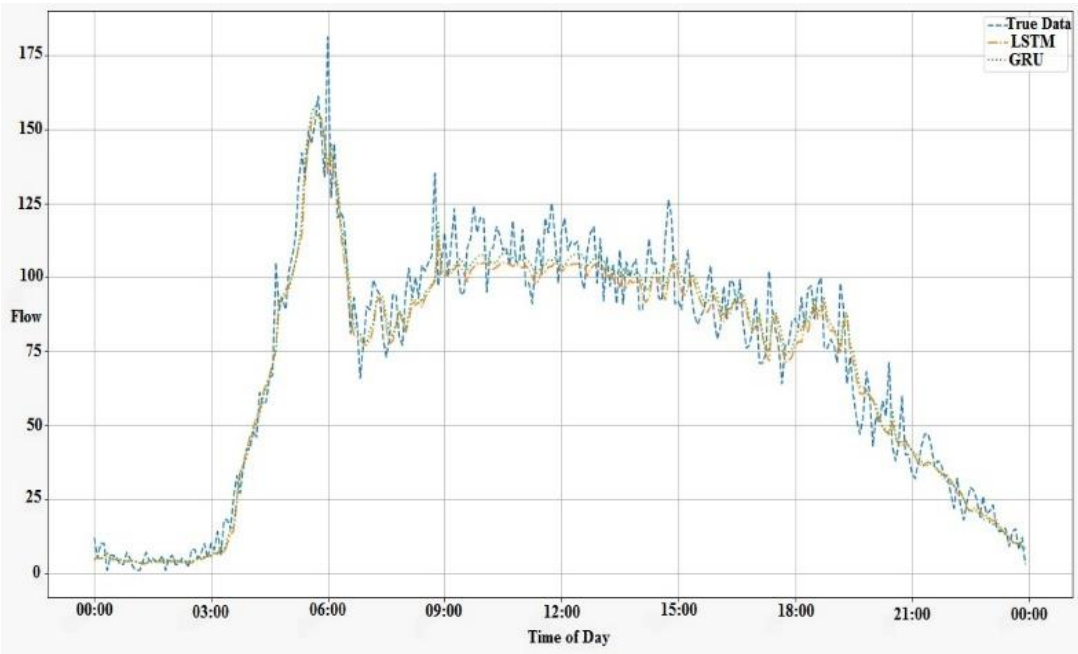

Figure 6. Performance comparison with the proposed models (LSTM and GRU) on PEMS dataset

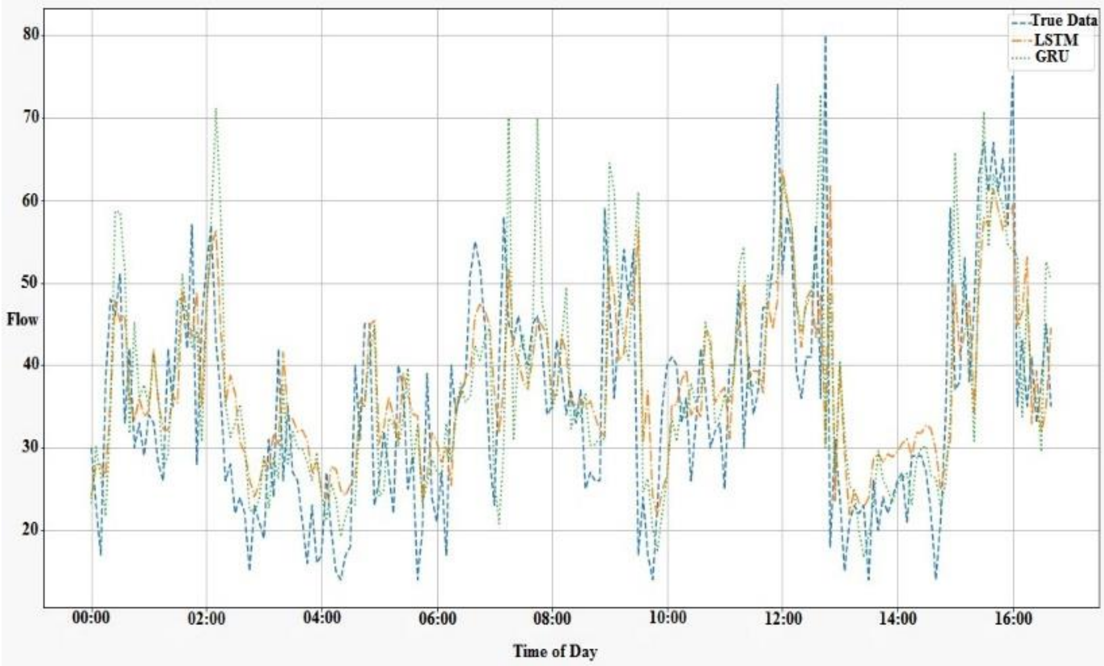

Figure 7. Performance comparison with the proposed models (LSTM and GRU) on TRANCOS dataset

Table 1. The results on PEMS dataset

Methods	Year	MAE	MAPE (%)	RMSE
STACKED AUTOENCODERS [30]	2014	34.1	——	50.0
ARIMA [46]	2016	19.17	——	29.0
LSTM NN [46]	2016	18.12	——	26.64
GRU NN [46]	2016	17.21	——	25.86
LSTM-M [48]	2018	32.2	——	47.04
DNN-BTF [36]	2018	19.12	——	27.91
LSTM [49]	2018	7.21	16.56	9.90
GRU [49]	2018	7.20	16.78	9.97
SAEs [49]	2018	7.06	17.80	9.60
Method [39]	2019	25.26	——	35.45
LSTM [50]	2019	20.59	——	30.86
PLSTM [50]	2019	19.98	——	30.21
PLSTM+ [50]	2019	19.06	——	28.78
MSTGCN [51]	2019	17.47	——	26.47
ASTGCN [51]	2019	16.73	——	25.27
Linear regression [6]	2019	7.60	——	10.32
Multi-layer perceptron [6]	2019	11.26	——	13.63
RBF network [6]	2019	20.58	——	27.10
RBF regressor [6]	2019	7.21	——	9.72
SMO reg [6]	2019	7.55	——	10.36
M2 [52]	2020	20.88	——	33.12
LSTM-BILSTM [53]	2021	12.63	——	16.72
Federated learning [54]	2021	7.96	——	11.04
MLP-NN [55]	2022	7.24	18.21	9.80
Proposed LSTM model Proposed GRU model	2022	7.14	16.37	9.74

Table 2. The results on TRANCOS dataset

Methods	Year	MAE	MAPE (%)	RMSE
Proposed LSTM model	2022	10.78	28.36	8.23
Proposed GRU model	2022	11.88	28.48	8.82

5. CONCLUSION

In this study, a short-term traffic flow prediction is introduced for ITSs to facilitate the scheduling of higher-level traffic that will enable travelers to save time and energy consumption. The current methods for predicting short-term traffic flow are incapable of effectively capturing the complex non-linearity of traffic flow, which leads to low prediction accuracy. To overcome this problem, this study introduces two novel models that deal with the non-linearity traffic data to predict traffic flow more efficiently. The first model uses an LSTM network followed by four dense layers, and the second model uses a GRU network followed by three dense layers. LSTM and GRU are used to extract the traffic flow temporal features, and the dense layers are used to perform the traffic flow prediction. The performance of the two models is evaluated using three metrics for the error analysis. The first metric is (MAE) which is used for continuous variable data. The second metric is (MAPE) which measures the accuracy of the prediction, typically shown as a percentage. The third metric is (RMSE) which is the standard deviation of residuals (prediction errors). The experimental results from MAE, MAPE, and RMSE indicate that the two models can properly capture the temporal correlations in the correlated traffic series and make reliable predictions. So, the two models can do this in specific cases and are able to suddenly capture trend changes. The proposed models provide improved predicting performance compared to the existing methods.

ACKNOWLEDGEMENTS

The authors would like to thank Fayoum University, Faculty of computers and Artificial Intelligence, Department of Computer Science, for supporting as of applied this work.

REFERENCES

[1] J. Zhang, F.-Y. Wang, K. Wang, W.-H. Lin, X. Xu, and C. Chen, “Data-driven intelligent transportation systems: a survey,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1624–1639, 2011, doi: 10.1109/TITS.2011.2158001.

[2] H. Zheng, F. Lin, X. Feng, and Y. Chen, “A hybrid deep learning model with attention-based conv-LSTM networks for short-term traffic flow prediction,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 11, pp. 6910–6920, 2021, doi: 10.1109/TITS.2020.2997352.

[3] I. Lana, J. D. Ser, M. Velez, and E. I. Vlahogianni, “Road traffic forecasting: recent advances and new challenges,” *IEEE Intelligent Transportation Systems Magazine*, vol. 10, no. 2, pp. 93–109, 2018, doi: 10.1109/MITS.2018.2806634.

[4] K. Zhang, L. H. Q. Tongyan, L. D. Wang, and C. H. Ruihua, “The latest achievements of Chinese national ITS architecture,” in *12th World Congress on Intelligent Transport SystemsITS AmericaITS JapanERTICO*, 2005.

[5] L. Li, S. Coskun, J. Wang, Y. Fan, F. Zhang, and R. Langari, “Velocity prediction based on vehicle lateral risk assessment and traffic flow: a brief review and application examples,” *Energies*, vol. 14, no. 12, pp. 5–34, 2021, doi: 10.3390/en14123431.

[6] J. Chang, J. Du, and H. Chung, “Vehicle traffic flow forecasting on caltrans PEMS dataset using machine learning algorithms and LSTM networks,” *Scientific and Practical Cyber Security Journal (SPCSJ)*, vol. 3, no. 3, pp. 25–39, 2019.

[7] J. Zheng and M. Huang, “Traffic flow forecast through time series analysis based on deep learning,” *IEEE Access*, vol. 8, pp. 82562–82570, 2020, doi: 10.1109/ACCESS.2020.2990738.

[8] M. Levin and Y.-D. Tsao, “On forecasting freeway occupancies and volumes (abridgment),” *Transportation Research Record*, no. 773, pp. 47–49, 1980.

[9] M. V. D. Voort, M. Dougherty, and S. Watson, “Combining kohonen maps with arima time series models to forecast traffic flow,” *Transportation Research Part C: Emerging Technologies*, vol. 4, no. 5, pp. 307–318, 1996, doi: 10.1016/S0968-090X(97)82903-8.

[10] B. M. Williams and L. A. Hoel, “Modeling and forecasting vehicular traffic flow as a seasonal arima process: theoretical basis and empirical results,” *Journal of Transportation Engineering*, vol. 129, no. 6, pp. 664–672, 2003, doi: 10.1061/(ASCE)0733-947X(2003)129:6(664).

[11] B. Ghosh, B. Basu, and M. O’Mahony, “Bayesian time-series model for short-term traffic flow forecasting,” *Journal of Transportation Engineering*, vol. 133, no. 3, pp. 180–189, 2007, doi: 10.1061/(ASCE)0733-947X(2007)133:3(180).

[12] P. Duan, G. Mao, W. Liang, and D. Zhang, “A unified spatio-temporal model for short-term traffic flow prediction,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 9, pp. 3212–3223, 2019, doi: 10.1109/TITS.2018.2873137.

[13] W.-C. Hong, Y. Dong, F. Zheng, and S. Y. Wei, “Hybrid evolutionary algorithms in a SVR traffic flow forecasting model,” *Applied Mathematics and Computation*, vol. 217, no. 15, pp. 6733–6747, 2011, doi: 10.1016/j.amc.2011.01.073.

[14] X. Feng, X. Ling, H. Zheng, Z. Chen, and Y. Xu, “Adaptive multi-kernel SVM with spatial–temporal correlation for short-term traffic flow prediction,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 6, pp. 2001–2013, 2019, doi: 10.1109/TITS.2018.2854913.

[15] G. A. Davis and N. L. Nihan, “Nonparametric regression and short-term freeway traffic forecasting,” *Journal of Transportation Engineering*, vol. 117, no. 2, pp. 178–188, 1991, doi: 10.1061/(ASCE)0733-947X(1991)117:2(178).

[16] H. Chang, Y. Lee, B. Yoon, and S. Baek, “Dynamic near-term traffic flow prediction: system-oriented approach based on past experiences,” *IET Intelligent Transport Systems*, vol. 6, no. 3, pp. 292–305, 2012, doi: 10.1049/iet-its.2011.0123.

[17] Y.-S. Jeong, Y.-J. Byon, M. M. C. -Neto, and S. M. Easa, “Supervised weighting-online learning algorithm for short-term traffic flow prediction,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, pp. 1700–1707, 2013, doi: 10.1109/TITS.2013.2267735.

[18] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, “Optimized and meta-optimized neural networks for short-term traffic flow prediction: a genetic approach,” *Transportation Research Part C: Emerging Technologies*, vol. 13, no. 3, pp. 211–234, 2005, doi: 10.1016/j.trc.2005.04.007.

[19] K. Y. Chan, T. S. Dillon, J. Singh, and E. Chang, “Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and levenberg–marquardt algorithm,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 644–654, 2012, doi: 10.1109/TITS.2011.2174051.

[20] W. Zheng, D. -H. Lee, and Q. Shi, “Short-term freeway traffic flow prediction: bayesian combined neural network approach,” *Journal of Transportation Engineering*, vol. 132, no. 2, pp. 114–121, 2006, doi: 10.1061/(ASCE)0733-947X(2006)132:2(114).

[21] M. Zhong, S. Sharma, and P. Lingras, “Short-term traffic prediction on different types of roads with genetically designed regression and time delay neural network models,” *Journal of Computing in Civil Engineering*, vol. 19, no. 1, pp. 94–103, 2005,

doi: 10.1061/(ASCE)0887-3801(2005)19:1(94).
[22] H. Yin, S. C. Wong, J. Xu, and C. K. Wong, "Urban traffic flow prediction using a fuzzy-neural approach," *Transportation Research Part C: Emerging Technologies*, vol. 10, no. 2, pp. 85–98, 2002, doi: 10.1016/S0968-090X(01)00004-3.
[23] M. -C. Tan, S. C. Wong, J. -M. Xu, Z. -R. Guan, and P. Zhang, "An aggregation approach to short-term traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 1, pp. 60–69, 2009, doi: 10.1109/TITS.2008.2011693.
[24] M. Cetin and G. Comert, "Short-term traffic flow prediction with regime switching models," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1965, no. 1, pp. 23–31, 2006, doi: 10.1177/0361198106196500103.
[25] L. Dimitriou, T. Tsekeris, and A. Stathopoulos, "Adaptive hybrid fuzzy rule-based system approach for modeling and predicting urban traffic flow," *Transportation Research Part C: Emerging Technologies*, vol. 16, no. 5, pp. 554–573, 2008, doi: 10.1016/j.trc.2007.11.003.
[26] P. Xie, T. Li, J. Liu, S. Du, X. Yang, and J. Zhang, "Urban flow prediction from spatiotemporal data using machine learning: a survey," *Information Fusion*, vol. 59, pp. 1–12, 2020, doi: 10.1016/j.inffus.2020.01.002.
[27] D. Zhang and M. R. Kabuka, "Combining weather condition data to predict traffic flow: a GRU-based deep learning approach," *IET Intelligent Transport Systems*, vol. 12, no. 7, pp. 578–585, 2018, doi: 10.1049/iet-its.2017.0313.
[28] M. Fouladgar, M. Parchami, R. Elmasri, and A. Ghaderi, "Scalable deep traffic flow neural networks for urban traffic congestion prediction," in *2017 International Joint Conference on Neural Networks (IJCNN)*, 2017, pp. 2251–2258, doi: 10.1109/IJCNN.2017.7966128.
[29] W. Huang, G. Song, H. Hong, and K. Xie, "Deep architecture for traffic flow prediction: deep belief networks with multitask learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 2191–2201, 2014, doi: 10.1109/TITS.2014.2311123.
[30] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, "Traffic flow prediction with big data: a deep learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 1–9, 2014, doi: 10.1109/TITS.2014.2345663.
[31] Z. Zhao, W. Chen, X. Wu, P. C. Y. Chen, and J. Liu, "LSTM network: a deep learning approach for short-term traffic forecast," *IET Intelligent Transport Systems*, vol. 11, no. 2, pp. 68–75, 2017, doi: 10.1049/iet-its.2016.0208.
[32] Y. Jia, J. Wu, and M. Xu, "Traffic flow prediction with rainfall impact using a deep learning method," *Journal of Advanced Transportation*, vol. 2017, pp. 1–10, 2017, doi: 10.1155/2017/6575947.
[33] N. G. Polson and V. O. Sokolov, "Deep learning for short-term traffic flow prediction," *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 1–17, 2017, doi: 10.1016/j.trc.2017.02.024.
[34] S. Du, T. Li, X. Gong, and S.-J. Horng, "A hybrid method for traffic flow forecasting using multimodal deep learning," *International Journal of Computational Intelligence Systems*, vol. 13, no. 1, pp. 1–12, 2020, doi: 10.2991/ijcis.d.200120.001.
[35] Y. Xiao and Y. Yin, "Hybrid LSTM neural network for short-term traffic flow prediction," *Information*, vol. 10, no. 3, p. 105, 2019, doi: 10.3390/info10030105.
[36] Y. Wu, H. Tan, L. Qin, B. Ran, and Z. Jiang, "A hybrid deep learning based traffic flow prediction method and its understanding," *Transportation Research Part C: Emerging Technologies*, vol. 90, pp. 166–180, 2018, doi: 10.1016/j.trc.2018.03.001.
[37] J. Bruna, W. Zaremba, A. Szlam, and Y. LeCun, "Spectral networks and deep locally connected networks on graphs," *Arxiv-Computer Science*, pp. 1–14, 2014.
[38] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting," *IJCAI International Joint Conference on Artificial Intelligence*, pp. 3634–3640, 2018, doi: 10.24963/ijcai.2018/505.
[39] W. Wei, H. Wu, and H. Ma, "An Autoencoder and LSTM-based traffic flow prediction method," *Sensors*, vol. 19, no. 13, pp. 1–16, 2019, doi: 10.3390/s19132946.
[40] P. Liu, Y. Zhang, D. Kong, and B. Yin, "Improved spatio-temporal residual networks for bus traffic flow prediction," *Applied Sciences*, vol. 9, no. 4, pp. 1–12, 2019, doi: 10.3390/app9040615.
[41] S. Zhang, Y. Yao, J. Hu, Y. Zhao, S. Li, and J. Hu, "Deep Autoencoder neural networks for short-term traffic congestion prediction of transportation networks," *Sensors*, vol. 19, no. 10, pp. 1–19, 2019, doi: 10.3390/s19102229.
[42] Caltrans, "Caltrans performance measurement system (PEMS)," *California Department of transportation*, 2019. https://pems.dot.ca.gov/. (acessed date: 01-Jan-2023)
[43] R. G. -G. -Olmedo, B. T. -Jiménez, R. L. -Sastre, S. M. -Bascón, and D. O. -Rubio, "Extremely overlapping vehicle counting," in *Pattern Recognition and Image Analysis*, Cham, 2015, pp. 423–431, doi: 10.1007/978-3-319-19390-8_48.
[44] B. K. Singh, K. Verma, and A. S. Thoke, "Investigations on impact of feature normalization techniques on classifier's performance in breast tumor classification," *International Journal of Computer Applications*, vol. 116, no. 19, pp. 11–15, 2015, doi: 10.5120/20443-2793.
[45] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
[46] R. Fu, Z. Zhang, and L. Li, "Using LSTM and GRU neural network methods for traffic flow prediction," in *2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, 2016, pp. 324–328, doi: 10.1109/YAC.2016.7804912.
[47] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, pp. 1724–1734, 2014, doi: 10.3115/v1/d14-1179.
[48] Y. Tian, K. Zhang, J. Li, X. Lin, and B. Yang, "LSTM-based traffic flow prediction with missing data," *Neurocomputing*, vol. 318, pp. 297–305, 2018, doi: 10.1016/j.neucom.2018.08.067.
[49] Xiaochus, "Traffic flow prediction," *GitHub*, 2018. https://github.com/xiaochus/TrafficFlowPrediction. (acessed date:21-Mar-2023)
[50] B. Yang, S. Sun, J. Li, X. Lin, and Y. Tian, "Traffic flow prediction using LSTM with feature enhancement," *Neurocomputing*, vol. 332, pp. 320–327, 2019, doi: 10.1016/j.neucom.2018.12.016.
[51] S. Guo, Y. Lin, N. Feng, C. Song, and H. Wan, "Attention based spatial-temporal graph convolutional networks for traffic flow forecasting," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 922–929, 2019, doi: 10.1609/aaai.v33i01.3301922.
[52] A. Boukerche and J. Wang, "A performance modeling and analysis of a novel vehicular traffic flow prediction system using a hybrid machine learning-based model," *Ad Hoc Networks*, vol. 106, pp. 1–10, 2020, doi: 10.1016/j.adhoc.2020.102224.
[53] C. Ma, G. Dai, and J. Zhou, "Short-term traffic flow prediction for urban road sections based on time series analysis and LSTM_BILSTM method," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 5615–5624, 2022, doi: 10.1109/TITS.2021.3055258.
[54] Y. Qi, M. S. Hossain, J. Nie, and X. Li, "Privacy-preserving blockchain-based federated learning for traffic flow prediction," *Future Generation Computer Systems*, vol. 117, pp. 328–337, 2021, doi: 10.1016/j.future.2020.12.003.

[55] A. N. -Espinoza *et al.*, "Traffic flow prediction for smart traffic lights using machine learning algorithms," *Technologies*, vol. 10, no. 1, pp. 1–11, 2022, doi: 10.3390/technologies10010005.

BIOGRAPHIES OF AUTHORS

Mohamed S. Sawah received an M.Sc. degree in Computer Science from Department Mathematics, Faculty of Science, Zagazig University, Egypt, in 2016. He can be contacted at email: me1900@fayoum.edu.eg.

Shereen Aly Taie is an Artificial Intelligence Professor in Computer Science Department, Faculty of Computers and Artificial Intelligence, Fayoum University, Egypt; She is the Vice Dean for Postgraduate Studies and Research, Faculty of Computers and Artificial Intelligence, and Head of Center of Electronic Courses Production, Fayoum University, Egypt. She held an M.S. in 2006 in Computer Science from Computer and Mathematics Department Faculty of Science, Cairo University, Egypt, and a Ph.D. in 2012 in Computer Science from Computer and Mathematics Department Faculty of Science, Cairo University, Egypt. She was awarded a B.Sc. degree in Computer Science by Faculty of Science, Cairo University in 1996. She was the Vice Dean for Education and Students Affairs, Faculty of Computers and Information, Fayoum University, Egypt. She is the head of "Artificial Intelligence Research Group" in Faculty of Computer and Information, Fayoum University, with overall 20 years of in-depth experience in teaching field for undergraduates and postgraduates along with researching experience gained during participation in +39 publications that had been issued in highly ranked international journals. My H-index is 14.03. Her teaching interest and research interest in Artificial Intelligence, Data Analysis, Data Science, Image Processing, Machine Learning, Algorithms, and Cloud Computing. She is a certified auditor in the National Authority for Quality Assurance and Accreditation, and a reviewer for research projects in the National Telecom Regulatory Authority (NTRA), Egypt. She can be contacted at email: sat00@fayoum.edu.eg.

Mohamed Hasan Ibrahim is an associate professor in Faculty of Computers and Artificial Intelligence, Fayoum University, Egypt, where he worked as a data scientist with deep and diverse domain experience. Mohamed is working as Vise Dean of Higher Studies and Research, Faculty of Computers and Artificial Intelligence. Mohamed has deep knowledge of data science and data analysis tools such as power Bi, Google Data studies, and IBM SPSS modeler. Mohamed is working as a lab manager at www.corelia.ai. He can be contacted at email: mhi11@fayoum.edu.eg and mib@corelia.ai.

Shereen A. Hussien is a Lecturer in Faculty of Computers and Artificial Intelligence, Fayoum University, Egypt where she has been a faculty member since 2018. she is graduated with a very good degree in B.Sc. from Faculty of Computer Science and Information Technology, Ain Shams University. from 2007-2017, she did her postgraduates M.SC. and Ph.D. from Faculty of Computers and Information, Helwan University. Her research interests related to computer vision and artificial intelligence merged with the new trend of IoT. she has experience in intelligent systems and adapting them in so many areas, including medical, agricultural, and educational fields. She carried out prediction and detection studies using various techniques. She can be contacted at email: sam26@fayoum.edu.eg.