Grenze International Journal of Engineering and Technology, January Issue

# A Novel Progressive Dual-Attention Residual Network with Bidirectional Long Short-Term Memory for Interpretable Traffic Flow Prediction

Shiju George[1], Asha Joseph[2], Shelly Shiju George[3] and Angel Shiju[4]
[1,4]Christ University, Bengaluru, Karnataka, India-560074
Email: drshijugeorge3@gmail.com
[2]New Horizon College of Engineering, Bengaluru, Karnataka, India
Email: asha.bti@gmail.com
[3]Alliance University, Bengaluru, Karnataka, India
Email: shellysgeorge@gmail.com

*Abstract*— **Traffic flow prediction constitutes a fundamental operational requirement for both Intelligent Transportation Systems (ITS) and urban planning systems. This research provides a detailed methodology integrating deep learning together with optimization features along with extraction methods and enhanced preprocessing techniques to enhance traffic flow prediction precision. An enhanced asymmetric Gaussian filtering method during preprocessing achieves successful reduction of noise while sustaining important patterns of traffic behavior. A VGG19-based Deep Sparse Auto encoder serves as a tool to extract features which makes it possible for deep learning to reveal complex traffic flow patterns. The prediction model implements its functionality through PDR-BiLSTM and its architectural components. Through its residual network structure and BiLSTM units, the model retains long-term dependencies in data patterns for stable predictions to be achieved in addition to utilizing dual-attention methods that help focus on essential spatial and temporal traffic characteristics. The Augmented Snake Optimizer uses parameter adjustments for model performance optimization to achieve prediction accuracy optimization of PDR-BiLSTM models. The experimental research indicates the proposed method achieves outstanding results by maintaining the lowest error rate with 99.5% accuracy and 99% precision recall F1-score.**

*Index Terms*— **Bidirectional Long Short Term Memory, Progressive Dual-attention Residual network, Asymmetrical cut down Gaussian filtering, VGG19-based Deep Sparse Auto encoder and Traffic Flow Prediction.**

## I. INTRODUCTION

Cities continuously spread outward from their centers to neighboring areas while facing worldwide challenges to their planning. A major effect of urban sprawl leads to elevated traffic levels in metropolitan areas [1]. The growing number of vehicles causes traffic congestion to become an urgent matter for urban planners who work throughout the world. The solution demands the creation of efficient traffic flow forecasting technologies that need special development for crowded urban roads and their heavily trafficked sections [2,3].

*Grenze ID: 01.GIJET.12.1.2*

The set of interacting elements within urban areas consists of various elements which include intersections together with pedestrian crossings and public transportation systems and multiple land uses [4,5]. Additional complexity elements between cities generate unpredictable traffic patterns which differ from standard highway traffic control systems. The changing dynamics of urban traffic cannot be properly tracked by forecasting methods which were originally developed for standard highways [6,7]. The development of innovative technologies and methodologies is crucial for improving traffic flow prediction in urban areas. Advanced data analytics, including Machine Learning and Artificial Intelligence tools, analyze real-time data collected from traffic cameras, GPS systems, and smart infrastructure sensors. These technologies extract valuable insights from traffic datasets, enabling the identification of hidden patterns, urban traffic flow trends, and anomalies, thereby enhancing the accuracy of predictions [8].

Urban planning strategies receive significant benefits when predictive modeling integrates with them to improve transportation infrastructure and reduce congestion. The implementation of adaptive traffic signal systems needs to be combined with public transit enhancements and promotions for cycling, walking and smart growth policies to fight urban sprawl [9,10]. The major contribution of this research is; An enhanced asymmetrical cut down Gaussian filtering method, noise reduction is improved while key aspects of the traffic pattern are preserved, providing a foundation for precise prediction. By utilizing a VGG19-based Deep Sparse Auto-encoder, complex traffic flow elements are captured and represented, and deep learning can be used to extract features. Long-term dependencies and spatial-temporal traffic patterns can be effectively captured by using a prediction model based on PDR-BiLSTM. Fine-tuning parameters and guaranteeing optimal prediction accuracy, the PDR-BiLSTM model's performance is optimized by the implementation of an Augmented Snake Optimizer.

## II. Literature Review

Several techniques were previously suggested in the literature related to traffic flow prediction among those some recent researches were reviewed in this section. Long-term traffic flow forecasting using a hybrid CNN-BiLSTM model has been proposed by Méndez M et al. [11] in 2023. The authorities now urgently require both long and short-term traffic flow forecasts due to the recent surge in road traffic in metropolitan cities. Making educated decisions about timely traffic congestions has become dependent on the availability of reliable traffic flow prediction techniques. Prior research had demonstrated that the accuracy of these techniques declined when long-term forecasts and urban traffic were taken into account. Du S. et al. [12] have suggested, a hybrid approach utilizing multimodal deep learning for traffic flow predicting in 2020. Forecasting traffic flow was thought to be a major issue for Intelligent Transportation Systems. Their approach's foundation module consisted of gated recurrent units (GRU) with an attention mechanism and one-dimensional CNN (1D CNN) because multi-modality traffic data is extremely nonlinear. While the later records the long-term temporal dependency, the former captured the features of local patterns.

Attention meets lengthy short-term memory a deep learning network for traffic flow forecasting, according to Fang W et al. [13] is recommended in 2022. Precise projection of forthcoming traffic patterns has been acknowledged as essential in numerous uses, forming a basic component of intelligent transportation networks. Due to the significant nonlinearity and volatility of traffic flow data, it is still difficult to provide accurate and timely traffic forecasts. Canonical long short-term memory (LSTM) networks were frequently prone to capture long-term dependencies in traffic flow evolution, but were less adept at focusing on minute-to-minute changes. An attention mechanism was added to LSTM networks for short-term traffic flow forecasting in order to address this difficulty.

Zheng J together with Huang M [14] utilized deep learning techniques to develop time series forecasting for traffic flow during 2020. The repeated issue of traffic congestion in major urban areas and their medium-sized counterparts puts sustainable development of cities at serious risk. Li et al. [15] compared the Intelligent traffic systems (ITS) prove to be among the most efficient solutions for tackling urban traffic congestion during the past few years. The correct prediction of traffic flow stood as a crucial factor for ITS operations. George & Santra [16] investigated the extensive length together with unpredictable patterns in traffic flow series exceeded the capabilities of existing traffic flow forecasting technologies. George & Santra [17] proposed time series analysis and deep learning (DL) to develop estimates for traffic flow but specifically addressed the existing problem. George & Santra [16] also presented a deep learning system designed for extended traffic flow estimation. Precise and dependable traffic flow prediction became essential for developing future management strategies since decision-makers need accurate forecasts which extend beyond short-term predictions like 24-hour predictions. Decision-makers encountered multiple complexities because traffic patterns were disordered along with being non-linear in structure.

## III. Proposed Methodology

In this research, an improved asymmetrical cut down Gaussian filtering is used in the preprocessing stage to efficiently reduce noise while maintaining important traffic pattern characteristics. A VGG19-based Deep Sparse Auto-encoder is utilized for feature extraction, utilizing deep learning to capture and represent complex traffic flow aspects. The prediction model is based on a PDR-BiLSTM architecture and a Progressive Dual-attention Residual network. Dual-attention mechanisms, which concentrate on pertinent temporal and spatial traffic aspects, are advantageous to this model. Long-term dependencies are captured and prediction robustness is enhanced by the residual network structure and BiLSTM units. An Augmented Snake Optimizer is used to maximize the PDR-BiLSTM model's performance. By effectively adjusting the model's parameters, this optimizer ensures the best possible prediction accuracy. Considering the aforesaid capabilities and objectives, the proposed method promises appreciable gains in traffic flow prediction through the combination of intricate preprocessing, feature extraction, cutting-edge neural network architecture, and optimization strategies.

### *A. Data Collection*

The input data are collected from Traffic Flow Forecasting Data Set. Globally, cities are experiencing an increase in traffic congestion. Growing urban populations, deteriorating infrastructure, haphazard and disorganized traffic signal timing, and a dearth of real-time data are all contributing factors.

### *B. An Improved Asymmetrical Cut down Gaussian Filtering based preprocessing*

The input data is initially given to an improved asymmetrical cut down Gaussian filtering based preprocessing approach. For preprocessing, Gaussian half-filters are dependable and easy to apply. According to Freeman and Adelson's research, the products of the important isotropic Gaussian by esteem to the x and y axes may be combined linearly to create the principal derivative of the 2D directional Gaussian $T_{\sigma,\vartheta}(a,b)$, guided at an angle $\vartheta$ given by Equation (1):

$$T_{\sigma,\vartheta}(a,b)=\cos(\vartheta)\bullet\frac{\partial T_{\sigma}}{\partial a}(a,b)+\sin(\vartheta)\bullet\frac{\partial T_{\sigma}}{\partial b}(a.b) \tag{1}$$

where $\sigma$ is the standard deviation of the Gaussian $T_{\sigma}$ and the coordinates of the pixel are written as $(a,b)$. As a result, the following is the truncated Gaussian derivative (Gd) that Nichani and Mehrotra proposed by Equation (2):

$$Gd_{\sigma}^{'}(a,b)=I(b)\bullet\left(1\frac{a}{2\pi\sigma^{4}}\bullet e^{\frac{-a^{2}+b^{2}}{2\sigma^{2}}}\right) \tag{2}$$

where, $I$ signifies the step function of Heaviside. Then the preprocessed data is given to VGG19-based Deep sparse auto-encoder technique for extracting relevant features.

### *C. VGG19-based Deep Sparse Auto-Encoder based Feature Extraction*

The pre-processed data is given as input to the VGG 19 architecture for extraction of relevant features. VGG-19 consists of 16 convolutional layers, 5 pooling layers, and 3 fully connected layers. It is valued for its simplicity, with sets of convolutional layers stacked to progressively increase the depth of the network. Max pooling layers are used in VGG-19 to reduce the spatial dimensions of the feature maps. There were two FC layers with 4096 neurons. Figure 1, shows the design of VGG19-based Deep sparse auto-encoder.

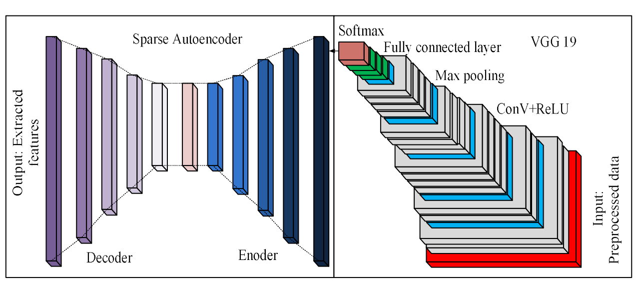


Figure 1. Architecture of VGG19-based Deep sparse auto-encoder

Deep sparse auto-encoder: An Autoencoder is designed to replicate its input at the output. Training a deep sparse Autoencoder involves two main steps. First, unsupervised greedy layer-wise pre-training is performed, where each layer extracts useful features from unlabeled data. Let $Encoder_{input}$ be the input to the encoder, which is transformed by an encoding function as shown in Equation 3.

$$Encoder_{input} = Fun(Dec_{input}) \tag{3}$$

The next step is to use Equation 2 to decode the encoded results ( $Encoder_{input}$ ) once the input data has been encoded by equation (4).

$$Decoder_{input} = Fun\left(Encoder_{input}\right) \tag{4}$$

Decoding's goal is to recreate the actual input. If the auto-encoder simply attempts to match its input to its outcome through the training phase, it might not be helpful since it will generate overfitting, which could yield subpar results on unseen samples.

*D. Progressive Dual-attention Residual network with Augmented snake BiLSTM (PDR-ABiLSTM)*

The extracted features are given to PDR-BiLSTM for accurate traffic flow prediction. The Progressive dual-attention residual network and BiLSTM framework are combined to form a PDR-BiLTSM framework.

*Progressive Dual-Attention Residual Network*

Rich location information about noteworthy items is often included in the first saliency prediction of the deepest layer of the backbone network. The aforementioned procedures should only be applied to three 3 x 3 dilated convolution branches because the features $F_T$ domain the inherent characteristics of the global semantic cues. In particular, encode channel-wise information first using equation (5) by relating global average pooling ( $GloAP$ ) to the input features ft in our attention unit.

$$R = U_{GloAP}(F_T) = \frac{1}{Y_{\cdot} \times O_{\cdot}} \sum_{y=1}^{Y} \sum_{o=1}^{O} F_T(y, o) \tag{5}$$

where $R$ The statistical vector RC represents the global features of the channel, whereas $Y_{\cdot}$ and $O_{\cdot}$ stand for the features' height and width, respectively. After that, use a straightforward gating mechanism made up of two successive nonlinear completely connected layers to capture channel dependencies. Additionally, the weight vector $Wei$ that is obtained shows how important each channel is, and it is calculated as equation (6):

$$Wei = Full_C\sigma\left(Full_C\delta\left(Wei\right)\right) = \sigma\left(S_2\left(\delta\left(S_1\left(Wei\right)\right)\right)\right) \tag{6}$$

where $S_1$ and $S_2$ denote the fully connected layer by the ReLU and sigmoid activation functions, correspondingly, and $Full_C\sigma$ and $Full_C\delta$ denote the associated layer. $Full_C\delta$ is set to 64 and $S_2$ is set to 2 in this experiment. Next, the channel-wise multiplication of the weight scalar $Wei$ and the input features $F_T$ of the attention unit yields the attention-enhanced feature mappings $F_T$ .

*BiLSTM*

BiLSTM type recurrent neural networks (RNNs) may derive dependencies from a data stream spanning both past and future contexts. For jobs that need sequence categorization, they are consequently particularly helpful. The first long short-term memory LSTM was developed to solve the problems of exploding and disappearing gradients that arise in traditional feed-forward neural networks.

The LSTM architecture contains of three gates: a forget gate, an output gate, and an input gate. It also has a memory cell that enables it to learn a series of long-distance dependencies and a hidden state. Two-way LSTM is a version of the conventional LSTM, which consists of two LSTMs executed concurrently on the input sequence, forward and backward.

Forward LSTM is used to acquire contextual information about the future, whereas reverse LSTM is used to capture contextual information about the past.

*E. Augmented Snake Optimizer*

Based on snake behaviour, the Augmented Snake Optimizer (ASO) [32] is a nature-inspired optimization algorithm. The objective of the fitness function is to minimize the weight parameters of the PDR-BiLSTM model. This minimization enhances the performance of PDR-BiLSTM, leading to higher accuracy. Table I presents the pseudocode of ASO, which has demonstrated substantial breakthroughs in the field of traffic flow prediction, yielding promising results.

TABLE I. PSEUDOCODE OF ASO ALGORITHM

**ASO Algorithm**

```
Initialize the binary Snakes randomly
while   do
   Evaluate each Snake in groups
   Find ideal male Snake malebest
   Find ideal female Snake femalebest
   Define Temperature
   Define food Quantity

   if   then
      do exploration
   elseif  then
      do exploitation
   elseif   then
      Snakes in Fight Mode
   else
      Snakes in Mating Mode
   end if
      Alter the worst male Snake and female Snake
   Compute the value of transfer function
   Update the Snakes

end while
Return bestSnake
```

A Progressive Dual-Attention Residual (PDR) network is integrated with a BiLSTM architecture in the proposed prediction model to enhance robustness and capture long-term dependencies. Model performance is further improved through dual-attention mechanisms that focus on relevant temporal and spatial traffic features. To optimize prediction accuracy, the parameters of the PDR-BiLSTM model are fine-tuned using an Augmented Snake Optimizer.

Overall, this comprehensive approach demonstrates significant advancements in traffic flow prediction, yielding promising results.

## IV. RESULT AND DISCUSSION

The experimental results and comments of the suggested method, which has been applied to the Python platform, are presented in this part. The suggested strategy is thoroughly assessed and contrasted with several other approaches to demonstrate its efficacy in terms of significant performance metrics like precision, MSE, accuracy, recall, and F1-score.

Table II, displays the experimental setup. The deep learning acceleration for this experiment was achieved using an NVIDIA Tesla V100 GPU on a high-performance computing system.

A Deep Sparse Auto encoder based on VGG19 was used for feature extraction to enhance the model's ability to detect spatial-temporal dependencies.

The PDR-BiLSTM architecture was trained for 100 epochs using an adaptive learning rate of 0.001 and a batch size of 64, optimized with the Augmented Snake Optimizer. The dataset was split with 80% for training and 20% for testing to ensure a balanced evaluation. The model's performance was assessed using key evaluation metrics, including accuracy, precision, recall, specificity, and F1-score, which confirmed the superiority of the proposed predictive strategy.

TABLE II.: EXPERIMENTAL SETUP

| Parameter | Specification |
|---|---|
| Hardware Platform | NVIDIA Tesla V100 GPU, 32 GB VRAM |
| Processor | Intel Xeon Silver 4216 @ 2.10 GHz |
| Memory | 128 GB RAM |
| Software Environment | Python , Tensor Flow |
| Training-Testing Split | 80% - 20% |
| Feature Extraction | VGG19-based Deep Sparse Autoencoder |

The Traffic Flow Forecasting Data Set is where the input data were gathered. Cities around the world are facing more traffic congestion. A lack of real-time data, haphazard and chaotic traffic signal scheduling, aging infrastructure, and growing metropolitan populations are all contributing issues.

*F. Experimental outcomes of introduced method*

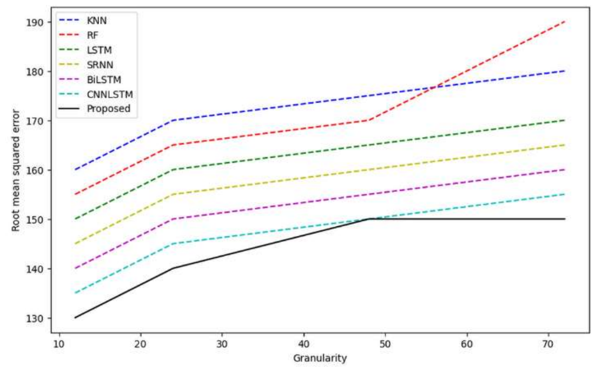


Figure 2. MSE analysis

Figure 2 shows the Mean Squared Error (MSE) analysis of the proposed approach. The granularity of the data i.e., how finely the time series is divided—can significantly affect forecast accuracy in traffic flow prediction studies. The data presented illustrates the Root Mean Square Error (RMSE) of various prediction models at 12, 24, 48, and 72 time units. Lower RMSE values indicate more accurate forecasts, as RMSE reflects the magnitude of prediction error. Among all the models, KNN consistently shows the poorest performance, while the proposed model achieves the best results, particularly at finer granularities. Notably, advanced models like CNN-LSTM and BiLSTM demonstrate significant improvements over traditional models such as KNN and RF, highlighting the effectiveness of deep learning methods in capturing complex traffic patterns. The proposed model outperforms both conventional models and other deep learning approaches.


Figure 3. Traffic data value analysis

Figure 3 illustrates the analysis of traffic data values using the proposed approach. To replicate actual traffic patterns, synthetic data has been generated for this traffic flow prediction study. The time_intervals array represents time steps ranging from 0 to 1400. To simulate real-world variability, traffic_flow is modeled as a sine wave with added random noise, capturing periodic fluctuations in the number of vehicles passing a specific point. Similarly, traffic_speed represents the average speed of vehicles, modeled using a cosine wave with additional random noise for increased realism. traffic_journey_time combines a sine wave with random noise and a phase shift to reflect the variability in travel times due to changing traffic conditions. These synthetic datasets enable the testing and validation of prediction models in a controlled yet realistic environment.

TABLE III: PERFORMANCE COMPARISON WITH EXISTING APPROACHES

| Techniques | Accuracy (%) | Precision (%) | Recall (%) | Specificity (%) | F1-score (%) |
|---|---|---|---|---|---|
| PDR-BiLSTM (Proposed) | 99.5 | 99.0 | 99.0 | 98.8 | 99.0 |
| CNN-BiLSTM | 97.8 | 96.5 | 96.2 | 96.0 | 96.3 |
| GRU | 96.2 | 95.0 | 94.7 | 94.5 | 94.8 |
| LSTM | 95.4 | 94.2 | 93.8 | 93.5 | 94.0 |
| W-CNN-LSTM | 94.8 | 93.5 | 93.0 | 92.7 | 93.2 |

The proposed PDR-BiLSTM model demonstrates exceptional performance in traffic flow prediction, outperforming traditional deep learning approaches during experimental evaluations. By integrating Progressive Dual- attention Residual (PDR) networks with a BiLSTM architecture, the model achieves 99.5% accuracy, along with 99% precision, recall, and F1-score—surpassing CNN-BiLSTM, GRU, LSTM, and W-CNN-LSTM, as shown in Table III. In addition to accelerating the learning process, the residual structure works in tandem with the dual-attention mechanism to optimize temporal and spatial feature selection. The parameters of the Augmented Snake Optimizer are finely tuned, leading to optimal prediction accuracy. Moreover, the VGG19-based Deep Sparse Auto encoder effectively extracts high-quality traffic features while preserving essential patterns through an asymmetrical cut-down Gaussian filtering technique. The PDR-BiLSTM model also achieves a high specificity of 98.8%, ensuring reliable traffic flow predictions by accurately identifying relevant patterns. Overall, the model proves its reliability and suitability for intelligent transportation systems by enhancing decision-making capabilities in urban traffic management.

GRU: Offers a balance between performance and computational efficiency. It achieves high accuracy and F1-scores with lower training and inference times, making it suitable for real-time traffic prediction applications. LSTM: Provides robust performance in capturing long-term dependencies in traffic data. However, it requires more computational resources and longer training times compared to GRU. PDR-BiLSTM, CNN-BiLSTM, and W-CNN-LSTM: Specific performance metrics for these models on traffic datasets are not readily available in the referenced studies. Further empirical evaluation would be necessary to assess their effectiveness in this domain. The performance observation of each method is depicted in TABLE IV.

TABLE IV. PERFORMANCE OBSERVATIONS

| Model | Architecture Summary | Strengths | Weaknesses | Ideal Use Cases |
|---|---|---|---|---|
| **PDR-BiLSTM** (Proposed) | Custom feature extraction (PDR) + BiLSTM | Tailored to domain, captures bidirectional dependencies | Complexity, needs domain knowledge | Specialized tasks (e.g., medical, signals) |
| **CNN-BiLSTM** | CNN for feature extraction → BiLSTM for sequence modeling | Captures spatial + temporal features | Requires large data, sensitive to CNN design | Audio, text, time-series with spatial traits |
| **GRU** | Recurrent unit with reset/update gates | Fewer parameters, faster training, effective | Slightly less expressive than LSTM | Time-series, low-resource tasks |
| **LSTM** | Classic RNN with forget/input/output gates | Handles long dependencies, robust | Slower than GRU, more parameters | General sequential tasks (text, signals) |
| **W-CNN-LSTM** | CNN with weights or sliding windows → LSTM | Focuses on relevant regions, balances context/locality | Needs careful tuning of weights/windows | Pattern recognition in signals, ECG, EEG |

## IV. CONCLUSION AND FUTURE SCOPE

The presented research offers an extensive new method for traffic flow estimation which stands essential for urban development planning and intelligent transportation systems. Resulting from deep learning and optimization together with feature extraction and advanced pre-processing techniques the proposed method brings substantial improvements to prediction accuracy levels. The employment of improved asymmetrical cut down Gaussian filtering results in noise reduction while sustaining traffic pattern essentials. The combination of

VGG19-based Deep Sparse Autoencoders for feature extraction enables detailed traffic pattern identification which is consolidated by PDR-BiLSTM as an architecture that constructs resilient long-term dependency prediction models. The implementation of dual-attention mechanisms enables the model to direct its attention toward essential traffic characteristics in time and space. The Augmented Snake Optimizer serves as an optimization method that completes parameter optimization for achieving the best possible model performance. The evaluation results exhibit high accuracy together with precision and recall performance as well as F1-score which proves the outstanding effectiveness of this proposed approach. This all-embracing method establishes exciting prospects for improving traffic flow prediction and enabling better urban traffic decisions through enhanced intelligent transportation systems development. The developed research framework delivers an effective method to predict traffic flow with precision in urban areas while making vital contributions to this field. Future research in this field should focus on developing dynamic adaption systems which work alongside real-time data feed mechanisms. Traffic flow projections become more dependable when real-time data analysis is used because urban traffic patterns keep shifting because of various events including accidents and road closures.

## REFERENCES


[1] Chen C, Liu Y, Chen L, Zhang C. Bidirectional spatial-temporal adaptive transformer for urban traffic flow forecasting. IEEE Transactions on Neural Networks and Learning Systems. 2022 Jun 30.

[2] Guo G, Yuan W. Short-term traffic speed forecasting based on graph attention temporal convolutional networks. Neurocomputing. 2020 Oct 14;410:387-93.

[3] Lee K, Eo M, Jung E, Yoon Y, Rhee W. Short-term traffic prediction with deep neural networks: A survey. IEEE Access. 2021 Apr 5;9:54739-56.

[4] Ma T, Antoniou C, Toledo T. Hybrid machine learning algorithm and statistical time series model for network-wide traffic forecast. Transportation Research Part C: Emerging Technologies. 2020 Feb 1;111:352-72.

[5] Zhu W, Wu J, Fu T, Wang J, Zhang J, Shangguan Q. Dynamic prediction of traffic incident duration on urban expressways: A deep learning approach based on LSTM and MLP. Journal of intelligent and connected vehicles. 2021;4(2):80-91.

[6] Lin X, Huang Y. Short-term high-speed traffic flow prediction based on ARIMA-GARCH-M model. Wireless Personal Communications. 2021 Apr;117(4):3421-30

[7] Lu H, Ge Z, Song Y, Jiang D, Zhou T, Qin J. A temporal-aware lstm enhanced by loss-switch mechanism for traffic flow forecasting. Neurocomputing. 2021 Feb 28;427:169-78.

[8] Ma Y, Zhang Z, Ihler A. Multi-lane short-term traffic forecasting with convolutional LSTM network. IEEE Access. 2020 Feb 17;8:34629-43.

[9] J Chen Z, Lu Z, Chen Q, Zhong H, Zhang Y, Xue J, Wu C. Spatial–temporal short-term traffic flow prediction model based on dynamical-learning graph convolution mechanism. Information Sciences. 2022 Sep 1;611:522-39.

[10] Osipov V, Nikiforov V, Zhukova N, Miloserdov D. Urban traffic flows forecasting by recurrent neural networks with spiral structures of layers. Neural Computing and Applications. 2020 Sep;32(18):14885-97.

[11] Méndez M, Merayo MG, Núñez M. Long-term traffic flow forecasting using a hybrid CNN-BiLSTM model. Engineering Applications of Artificial Intelligence. 2023 May 1;121:106041.

[12] Du S, Li T, Gong X, Horng SJ. A hybrid method for traffic flow forecasting using multimodal deep learning. International journal of computational intelligence systems. 2020 Jan;13(1):85-97.

[13] Fang W, Zhuo W, Yan J, Song Y, Jiang D, Zhou T. Attention meets long short-term memory: A deep learning network for traffic flow forecasting. Physica A: Statistical Mechanics and its Applications. 2022 Feb 1;587:126485.

[14] Zheng J, Huang M. Traffic flow forecast through time series analysis based on deep learning. IEEE Access. 2020 Apr 27;8:82562-70.

[15] Li Y, Chai S, Ma Z, Wang G. A hybrid deep learning framework for long-term traffic flow prediction. IEEE Access. 2021 Jan 11;9:11264-71

[16] George, S., Santra, A.K. Traffic Prediction Using Multifaceted Techniques: A Survey. Wireless Pers Commun 115, 1047–1106 (2020). https://doi.org/10.1007/s11277-020-07612-8

[17] George, S., & Santra, A. K. An improved long short-term memory networks with Takagi-Sugeno fuzzy for traffic speed prediction considering abnormal traffic situation. Computational Intelligence, 36(3), 964-993, 2020 https://doi.org/10.1111/coin.12291

[18] George, S., & Santra, A. K. (2020). Fuzzy Inspired Deep Belief Network for the Traffic Flow Prediction in Intelligent Transportation System Using Flow Strength Indicators. Big Data, 8(4), 291-307, 2020


.