# A Secure Intelligent System for Internet of Vehicles: Case Study on Traffic Forecasting

Youcef Djenouri, Asma Belhadi, Djamel Djenouri, Gautam Srivastava, and Jerry Chun-Wei Lin*

***Abstract*—While significant efforts have been made for vehicle-to-vehicle communications, which now enable the Internet of Vehicles (IoV). Current IoV solutions are unable to capture traffic data both accurately and securely. Another drawback of current IoV models based on deep learning is the methods they use to tune hyperparameters. In this paper, a new system called secure and intelligent system for the Internet of Vehicles (SISIV) is developed. A deep learning architecture based on graph convolutional networks and an attention mechanism are used. In addition, blockchain technology is used to protect the data transmission between nodes in the IoV system. Moreover, the hyperparameters of the generated deep learning model are intelligently selected using a branch-and-bound technique. To validate SISIV, experiments were conducted on four networked vehicle databases dealing with prediction problems. In terms of forecasting rate ($>$ 90%), F-measure ($>$ 80%), and attack detection ($<$ 75%), the results clearly showed the superiority of SISIV over the baseline systems. Moreover, compared to state-of-the-art solutions based on traffic prediction, SISIV enables efficient and reliable prediction of traffic flow in an IoV context.**

***Index Terms*—Deep Learning, Internet of Vehicles, Blockchain, Graph Convolution Network.**

## I. Introduction

The Internet of Things (IoT), wireless networking, big data as well as artificial intelligence have propelled research to new heights in many facets of our lives [1]–[3], as well as many application domains such as human behaviors [4], smart privacy [5]–[7], smart homes [8], smart transportation [9]–[11] as well as Internet of Vehicles (IoV) [12]–[14]. IoV is regarded as one of the most innovative technologies of the modern era as well as the evolution of our societies [15], [16]. AI plays an important role in the modernization of vehicle technology [17]–[19], as well as several pieces of research have been conducted in this direction. Wang et al. [20] built a two-level aided vehicular network framework around federated learning. They created a new federated learning participant decision mechanism that is supported by mobility as well as reduced the cost of federated learning with a distributed joint resource allocation strategy. Li et al. [21] proposed an emergency information dissemination approach based on the social IoV to create inter-vehicle social ties without human interaction as well as exchange emergency information through stable vehicle-to-vehicle links. The information diffusion problem was reformulated as an influence maximization problem based on a vehicle link graph. They devised a social IoV-based emergency information influence maximization method to increase the influence range by picking some influential seed vehicles as well as raising the influence of others. Liu et al. [22] suggested a multi-unmanned aerial vehicle enabled mobile IoV paradigm in which unmanned aerial vehicles track to serve mobile vehicles as well as deliver downlink information to them during flight. The system throughput is maximized by concurrently optimizing vehicle communication scheduling, unmanned aerial vehicle power allocation, as well as unmanned aerial vehicle trajectory, taking into account the limits of anti-collision as well as communication interference between the unmanned aerial vehicles. The non-convex optimization problem is broken down into three subproblems, i) communication scheduling optimization, ii) power allocation optimization as well as iii) unmanned aerial vehicle trajectory optimization. These subproblems may be able to be handled via successive convex approximation. To find the best solution, a combined iterative optimization approach for the three subproblems is proposed.

All the aforementioned solutions suffer from many issues, which may be able to be highlighted in the following:

1) They do not provide a framework that guarantees secure transfer of data in the IoV network.
2) Missing accurate deep learning models that are able to learn from the different features of the IoV network.

To address these issues, this study proposes SISIV; a secure as well as intelligent system connected to vehicles in the context of the IoT. Our contributions are as follows:

1) We propose a novel framework, named SISIV (**S**ecure and **I**ntelligent **S**ystem for **I**nternet of **V**ehicles), which consider both privacy preservation as well as learning issues. Privacy is protected by implementing an efficient blockchain-based technique to secure data transfer in the IoV, and learning is protected by a graph neural network that learns from the visual features of the IoV. We also develop the graph attention network to improve the learning process by focusing more on the relevant visual features.
2) We present a branch-and-bound optimization strategy to optimize the hyperparameters of the deep learning

Y. Djenouri is with the SINTEF Digital, Forskningsveien 1, 0314, Oslo, Norway.

A. Belhadi is with with the Department of Technology, Kristiania University College, Oslo, Norway.

D. Djenouri is with CSRC, Dep. of Computer Science & Creative Technologies, University of the West of England, Bristol, UK.

G. Srivastava is with the Department of Mathematics & Computer Science, Brandon University, Brandon, Canada as well as with the Research Centre for Interneural Computing, China Medical University, Taichung, Taiwan

J. C. W Lin is with the Department of Computing, Mathematics, and Physics, Western Norway University of Applied Sciences, Bergen, Norway. (*Corresponding author)

architecture used in SISIV. Rather than using exhaustive search methods, the strategy considers the hyperparameter space and uses a heuristic to intelligently explore the enumeration tree.

3) On four different connected vehicle datasets dedicated to forecasting problems, we show that the SISIV framework gives promising results compared to the baseline IoV solutions in terms of forecasting rate, runtime, as well as detected attacks.

The rest of the paper is organized as follows. Section II gives a review of the literature related to security in IoT as well as IoV applications. Section III presents a detailed explanation of the SISIV framework. A performance evaluation of the SISIV framework is provided in Section IV. Section V shows the open research scope for IoV. Section IV draws the conclusion.

## II. Related Work

This section reviews the literature on security in IoT as well as IoV applications.

### A. Security in IoT

For detecting infiltration in IoT devices, Nie et al. [23] created a generative adversarial network. The features were chosen first in order to handle the sensor data appropriately. A single attack was then discovered using the generative adversarial network. To identify as well as comprehend the behaviours of various attacks, a combination of several intrusion detection architectures was deployed. Wang et al. [24] devised a method for analyzing the trust in mobile edge nodes in order to improve IoT device reliability as well as mitigate network assaults. It is a graph-based model in which sensors are vertices as well as point-to-point connections are edges. Djikstra algorithm calculates the sensor node's trust after measuring as well as improving the individual sensor node's trust score. Nagarajan et al. [25] investigated the capability of gateway nodes for gathering as well as safeguarding data for IoT applications. A deep learning technique was used in fog systems to study as well as train the acquired data. The proposed technique not only learns from IoT data but also takes into account relief formulae to deal with the difficult restrictions of the sensitive data acquired. Belhadi et al. [4] created a fusion model to detect anomalies in group trajectories from pedestrian collective behavior data. This model was developed in the context of intelligent transportation. Several data mining as well as deep learning-based systems were created, as well as group anomalies were determined using solutions-based neighbourhood computation as well as clustering. Zekry et al. [26] suggested two convolution LSTM deep learning models to detect anomalous data from IoT sensors as well as avert cyber-attacks. Aloqaily et al. [27] proposed a solution for intrusion detection as well as prevention in IoT based on clustering, in which cluster heads are chosen so that services as well as providers may be able to communicate with third-party entities. The authors employed a decision tree to choose the attributes as well as classify the attacks after using a deep belief function to minimize data dimensionality as well as discover positive trustworthy service requests.

### B. IoV as well as their Related Applications

Several research efforts were dedicated to vehicle-to-vehicle communications [28], vehicular ad hoc networks [29], [30], as well as related applications, e.g., [31], [32]. This evolved to the emergence of IoV in which vehicles are interconnected as well as connected to the internet. Abdellatif et al. [33] created an active learning framework that reacts to unusual road scenarios using data collected from onboard sensors as well as other vehicles. The framework looked at three different ways that vehicles get information, i.e., by sharing labels, data, or a combination of the two. Deep learning models could be utilized by parked vehicles (PVs). Li et al [34] modelled the time of arrival as well as the duration of parking using the Weibull as well as dual Gamma distributions. PVs were also persuaded to share their unused compute assets through a contract-based incentive mechanism. Xing et al. [35] looked into the relationship between connected vehicle energy consumption as well as driving styles They investigated the impact of the amount of energy consumed on the accuracy of driver behavior detection as well as motion/trajectory prediction systems. A deep learning-based approach was used for time-series modeling. The results showed that anticipating driving behaviours as well as accurately predicting vehicle motion is difficult for vehicles with high energy usage. RNN-LF (Recurrent Neural Network for Long-term Flows) was created by Belhadi et al. [36] to anticipate long-term traffic data represented by flow distribution. It drew on a variety of data sources as well as contextual knowledge, such as weather data. Xu et al. [37] proposed TripRes, a traffic flow prediction system that relied on a city map. In this system, the collection of large regions is first selected, as well as the deep spatiotemporal residual network is then trained to learn from the current traffic condition as well as infer future traffic flow predictions for similar regions. Peng et al. [38] proposed a hybrid-based model for long-term traffic flow prediction, called GCN-LSTM. Deep learning concepts including graph convolution neural network (GCN) as well as LSTM were combined in this algorithm. GCN learns traffic data's spatial patterns, whereas LSTM learns traffic data's temporal patterns. Xu et al. [39] created an edge-based system for IoV. The residual network is used to learn the future services of the IoT system. Multi-objective optimization ass also used to reduce the overall system's time as well as energy costs. An intelligent framework that can process various data from connected cars has been developed by Sun et al. [40]. The use of adaptive data cleaning allowed for the elimination of noise, thus improving the data collection process. The method uses an autoencoder with large short-term memory and has four layers for training the data cleaning mechanism. An insightful paradigm for analyzing heterogeneous IoV data was developed by the authors. The IoV data were cleaned using an adaptive data cleaning method based on autoencoder long-term memory, which had four layers. This helped to minimize the amount of noise in the data.

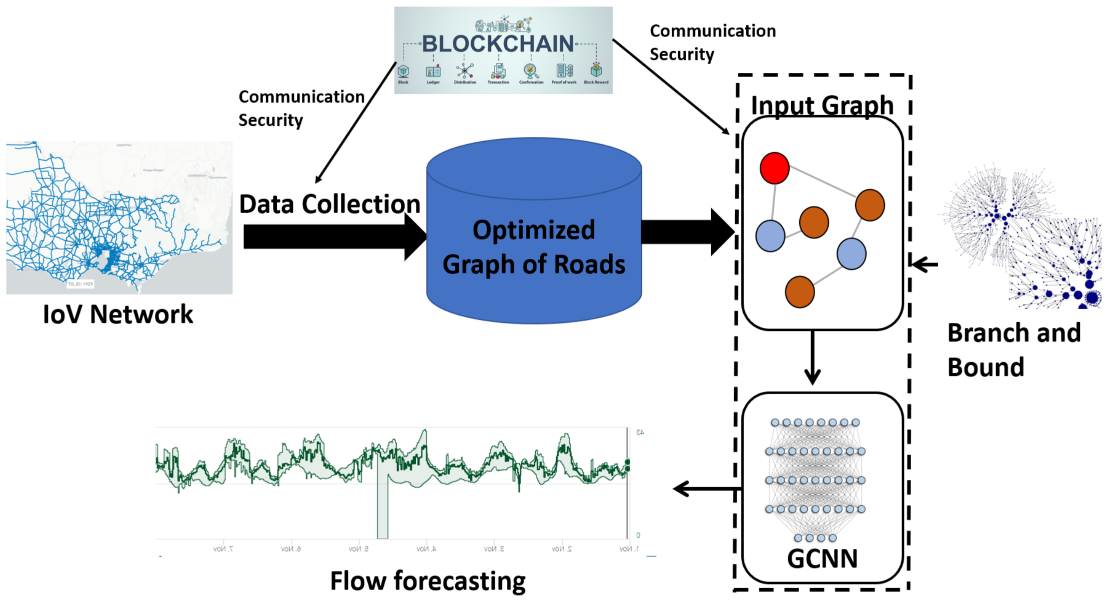


Fig. 1. SISIV Framework: The optimized graph of the roads is first created from the raw IoV data. The graph convolutional network is then trained for traffic forecasting while the branch-and-bound method is used to determine the optimal parameters of the trained model. Blockchain technology is used to secure the data communication.

### C. Discussions

This literature review reveals that IoV technologies have a number of flaws. First, they are unable to securely manage sensitive IoV data collection. The accessible IoV information would greatly aid in improving the model's accuracy. The second difficulty is hyper-parameter optimization, which necessitates the adjustment as well as tuning of numerous parameters during the training phase. To solve the deal with these problems, we provide a new secure as well as intelligent framework for IoV in the following section.

## III. Designed Secure and Intelligent System for Internet of Vehicles Framework

### A. Principle

Figure 1 depicts the framework of the proposed solution that is based on Branch-and-Bound for a smart hyper-parameters tuning, as well as graph convolution neural network (GCN) for IoV data handling. We also create a secure-based system for transferring data securely. First, the IoV data from the sensors is retrieved. After the hyper-parameters have been tweaked, deep learning is used to determine the forecasting of IoV. We present traffic data collected by IoT sensors as graphs to obtain the geographic structure inherent
on the work of Gue et al. [41], we propo
graph implementation by adopting the invert
which is in contrast to the usual graph
the road network where the road segment
as edges and the intersections as nodes.
road segments are considered as graph no
connections between adjacent pairs of ro
edges of these nodes are formed. In th
the GCN with attention mechanism and the branch-and-bound method are applied to learn from the optimized graph. In particular, the attention mechanism aims to capture the relevant features from the graph data, and the branch-and-bound method aims to determine the optimal parameters of the GCN model. The IoV system uses blockchain technology to secure the communication between the different nodes. In particular, this ensures the confidentiality of the data collected during training and also ensures that the trained model is secured against unexpected changes for the deployment phase. The following sections provide descriptions of the SISIV components.

### B. Graph Convolution Neural Network (GCN)

The appropriate management of spatial data represents the principal challenge for traffic forecasting. The structure of the urban road network is becoming increasingly complex as transportation technology develops at a rapid pace. Traditional CNNs are incapable of meeting today's demands. GCNs [42] have been experimentally demonstrated suitable for traffic forecasting. It may be able to fully capture the spatial properties of traffic data, improving the model's overall prediction performance. The GCN model is employed in this
dden layers, with the ReLU function as
n function. There is only one parameter
kernel. The number of parameters is
when each and every convolution kernel
ble parameter. However, studies have
a significant number of parameters
impact on the model performance. To
nvolution may be able to be defined as
hboring nodes employing an attention

technique. An attention mechanism is a piece of software that forces the model to concentrate on learning as well as absorbing crucial data. The primary strategy of the attention mechanism is to include it in the GCN model. Previous research reveals that the model may be able to perform parallel computing across nodes as well as overcome spatial convolution limitations. It also has the ability to learn inductively. A graph attention mechanism must first be used before the degree of linkage between nodes may be able to be calculated. In this mechanism, the graph attention layer's input is a node feature. Without processing, the attention coefficient between nodes is quite complicated. To normalize the attention coefficients of each and every node, we use the softmax function. The normalized attention coefficient is then utilized to differently aggregate the information of surrounding nodes. We create here a learnable function based on the attention mechanism to acquire the relationships between adjacent nodes, i.e., the local graph structure.

### *C. Branch and Bound*

The Branch-and-Bound algorithm has shown its efficiency for discrete as well as combinatorial optimization problems, as well as mathematical optimization [43], where the lowest bound yet discovered are tracked, compared with the possible solutions, as well as then only keeps a possible solution that is inferior to the lowest bound yet discovered as the new lowest bound. This may be able to solve only minimization problems. But this does not represent a drawback as any maximization problem may be able to be transformed into a minimization one by multiplying the objective function by $-1$. Convex problems are the only ones for which the global optimum is guaranteed. Forming a rooted tree of viable solutions to the problem is what branching is all about. Then it is possible either to conduct an extensive search (examine all of the tree's branches) or eliminate searching through some branches (pruning) that are know not to include solutions. This applies to convex problems in one of the following scenarios.

1) The value of a variable (a constraint), in this example the value of a hyper-parameter, is infeasible. As a result, we remove all branches that are tied to that value. Because going down merely adds more limits, there's no need to go any longer if and only if one is already impossible.
2) If the estimated best solution through the explored branches is worse than the current one, then the exploration phase of the current branch is terminated then move forward to others.
3) A solution is discovered as well as no better one may be able to be discovered by moving further down the branch (as this will only add more constraints). In this case, all that may be able to be done is for a comparison of this solution to the best found so far.

Notice that if and only if the problem is not convex, any further anticipated pruning may cause the missing of a local or global optimum.

To clarify which variable should be branched, we use an illustrative example of a binary problem. A binary problem is an optimization problem in which the variables are the interval [0,1]. Thus, each and every variable $x_i$ for $i = \{1, 2, 3, \dots, n\}$ is represented in this given form:

$$X \in \{0, 1\}^n, \tag{1}$$

where $n$ is the number of variables in the problem. This may be able to be relaxed to:

$$0 \leq x_i \leq 1, for : i = \{1, 2, 3, \dots, n\} \tag{2}$$

For instance, consider this solution:

$$X = [0.1, 0.6, 1, 0, 1] \tag{3}$$

For exploration, we choose a branch on the variable whose value is closest to 0.5 (in this case $x_2$). We propose two exploration strategies:

1) **Depth first strategy:** It takes a branch down to the bottom until reaching a point we cannot go any farther, then works the way back up.
2) **Breadth first strategy:** Various branches are explored at each and every depth, then continue deeper as well as repeat the process until reaching the bottom.

The depth-first method is theoretically the better method because it leads to a solution faster by imposing more and more constraints and it can be compared with other solutions, which speeds up pruning.

### *D. Blockchain*

We secure the proposed framework with blockchain technology. To set up a secure decentralized traffic forecasting system, we are developing a dedicated consortium blockchain. To make the system more functional, we select a certain number of Road Side Units (RSUs) as approved miners. We then update the hardware configuration of the RSUs for this purpose and to provide robust computational, storage, and networking capabilities for evaluating local model updates transmitted from remote vehicles. In this way, both inaccurate and unreliable updates can be detected. They use carefully tuned consensus procedures to create a new block of records of qualified local model updates. An iteration of the global model training for implementing predictions includes the following steps after installing the consortium blockchain:

1) Local algorithm execution: Each and every vehicle first runs the forecasting algorithm on its local dataset. This permits local outputs to be generated as well as relayed to the nearest miner.
2) Output control: Miners are hired to receive as well as verify local outputs in order to get defined token rewards. A new data block is created by the miner whenever an efficient approach is used to filter out both spurious and low-quality local outputs. All local outputs that meet the qualification criteria are stored in this new data block.
3) Merging: The consortium blockchain logs the new block, which includes the most recent local outputs, as well as instructs all participants to download the most recent block data. Each and every person may be able to compute the global outputs by knowing the local outputs of the other participants.

We optimize the algorithm created in [44] to accurately implement the above steps. Each individual miner is tasked with downloading a standardized test dataset and determining whether or not the vehicle's local model updates are qualified. Depending on the particular accuracy requirements of the algorithm, each miner employs specific filtering algorithms to screen out hostile entities with bugs or poisoning attacks. We use the consensus method to analyze and filter harmful entities to prevent them from affecting normal system operation. This allows us to identify and reduce the harmful impact of hostile entities on the consortium blockchain. We use the flexibility of the consensus algorithm to protect against future security threats. Only the appropriate local results are combined to obtain the latest global result. In this way, low-quality local results can be eliminated and accurate prediction can be achieved. Unlike previous consensus algorithms such as computationally intensive proof-of-work, communication-intensive Byzantine fault tolerance, and unfair proof-of-stake, the new method is based on practical Byzantine fault tolerance and allows flexible mining without a fixed miner group. This is achieved through the practical application of Byzantine fault tolerance.

**Algorithm 1** SISIV Algorithm

```
1: Input: IoV: Raw IoV data;
   V = {V_1, V_2..., V_n}: The set of n IoV training data;
2: Output: Forecast_IoV: The set of predicted traffic;
3: V ← CollectionFromSensors(IoV);
4: G ← ConstructOptimzedGraph(V);
5: R_IoV ← ∅;
6: while Blockchain(G) is secured do
7:    for G_i ∈ G do
8:       Forecast_i ← BB(GNN(G_i));
9:       Forecast_IoV ← Forecast_IoV ∪ Forecast_i;
10:   end for
11: end while
12: return Forecast_IoV.
```

### E. The SISIV Algorithm

Algorithm 1 presents the pseudocode of the SISIV algorithm. It starts by collecting and constructing the training data in (lines 3 and 4). The data is collected from a set of sensors and an optimized graph is created from the collected data. The whole dataset $G$ is first parsed data instance by data instance, and then the GCN is performed with attention mechanism. The hyper-parameters are also optimized using the branch-and-bound method (lines 7 to 10). The process is secured by blockchain technology (lines 6 to 11). The predicted results are returned (line 12).

## IV. Performance Evaluation

SISIV is compared to state-of-the-art IoV solutions in this section. The evaluation is based on four standards benchmark datasets for connected vehicles [1]. The datasets include labelled data with ground truth, as well as the result of all datasets are averaged for comparison metrics as well as presented in the following. The following is a brief description of these benchmarks:

1) **Astyx**: This IoV dataset provides high-resolution radar data as well as 3D object detection using radar, LIDAR, and camera data. It contains $546$ frames and is over $350MB$ in size.
2) **Deep Drive**: The dataset includes over $100,000$ of video sequences with various annotations, including image-level labels, object bounding boxes, drivable areas, lane markers, and segmentation of full-frame instances. The dataset is geographically, ecologically, and weather diverse.
3) **Landmarks**: An open-source Google database for detecting artificial and natural landmarks. It was used in the 2018 Kaggle competitions for landmark detection and retrieval of 2 million images, including $30,000$ in interesting locations from around the world.
4) **Accidents**: A nationwide U.S. database of motor vehicle crashes covering $49$ states from February 2016 through December 2020. It uses APIs from state departments of transportation, law enforcement, traffic cameras, and sensors to offer live traffic data. Three million accidents have been reported.

The forecasting rate is calculated as well as used as a comparative statistic to evaluate the proposed framework. It is defined as follows:

$$FR = \frac{CF}{|T|} \times 100, \tag{4}$$

where $CF$ is the number of the correctly forecasted test samples, as well as $T$ is the size of the test dataset.

The evaluation is also performed using different measures: precision (P), recall (R), and F-measure (F), which are defined as follows:

$$P = \frac{TP}{TP + FP} \tag{5}$$

$$R = \frac{TP}{TP + FN} \tag{6}$$

$$F = \frac{2 \times P \times R}{P + R} \tag{7}$$

where $TP$ denotes the number of samples whose true label and predicted label are both positive. $FP$ denotes the number of samples with negative true label and positive predicted label. $FN$ denotes the number of samples with positive true label and negative predicted label. Note that these measures are common measures for evaluating traffic forecasting methods. The empirical testing was carried out on a machine with a 64-bit core i7 processor, Windows 10, as well as 16 GB of RAM. The CPU host is a 2.27 GHz Intel Xeon E5520 quad-core 64-bit processor. The GPU is an NVIDIA Tesla C2075, which has $448$ CUDA cores ($14$ multiprocessors with $32$ cores each) as well as runs at $1.15$ GHz. It has a global memory of $2.8$ GB, a shared memory of $49.15$ KB, as well as a warp size of $32$. Single precision is used on both the CPU as well

[1]https://www.kaggle.com/datasets

as GPU. The SISIV framework is compared to the following baseline solutions:

1) **RNN-LF (Recurrent Neural Network for Long-term Flows)** [36]: It is a recurrent neural network developed with the goal of predicting long-term traffic data represented by the distribution of traffic flows. It uses a variety of data sources as well as contextual knowledge, such as information about traffic and weather.
2) **TripRes** [37]: It uses the city map to efficiently estimate traffic flow. First, a collection of large areas is defined. Then, a deep spatiotemporal residual network is trained to learn from the current traffic scenario and predict future traffic flows of comparable regions based on what it learned from the previous situation.
3) **GCN-LSTM** [38]: It uses a hybrid architecture for accurate prediction of long-term traffic flow. It uses graph CNN in addition to LSTM. LSTM is responsible for learning the temporal patterns of the traffic data, while GCN is responsible for learning the spatial characteristics of the traffic data.
4) **Gra-TF** [45]: It is a graph-level forecasting approach used to develop an integrated as well as improved forecasting model using ensemble learning. Several strategies are used in this design model to reduce uncertainty in IoV systems.

### A. *Parameters Setting*

In this section, the results of the parameter setting of the SISIV framework are explained. In this work, the branch-and-bound optimization approach was used to optimize the hyper-parameters of the deep learning model. We applied both the deep-first and breadth-first methods, and the best value from both was returned. The number of epochs can be set between 100 and $1,000$, the learning rate between 0 and 1, and the number of batches between 16 and 512. For each of the four benchmark datasets, the branch-and-bound technique examined the space of hyperparameters and found the optimal parameters for our model in terms of forecasting rate. Table I summarizes the values of these parameters.

TABLE I
BEST PARAMETERS OF SISIV.

| Dataset | epochs | learning rate | batches |
|---|---|---|---|
| Astyx | 257 | 0.43 | 64 |
| Deep Drive | 315 | 0.49 | 32 |
| Landmarks | 439 | 0.55 | 64 |
| Accidents | 544 | 0.82 | 32 |

### B. *Experimental Results*

*1) Forecasting Rate:* For the four datasets above, in the initial tests, we compare the forecasting rate of SISIV with that of the baseline solutions by varying the number of traffic data to be predicted as input from 50 to 500 in the test set. Figure 2 shows that SISIV outperforms the four baseline algorithms. SISIV's forecasting rate reached 95% when processing 500 of traffic data from the Deep Drive dataset, while the rates of the other models were below 80%. These results were obtained by combining a GCN with an attention mechanism that takes advantage of the different information propagations as well as the features injected into the generated GCN. This enables more accurate observation predictions and helps in developing smarter IoV decisions. In addition, the branch-and-bound technique can effectively adjust the hyperparameters of the various deep learning models used in SISIV to achieve the highest possible forecasting rate.

*2) Runtime:* The second experiment compares the training runtime of SISIV with the baseline methods, with the error loss value set to $0.005$. When increasing the amount of traffic input from $5,000$ to $50,000$, SISIV outperforms the four baseline models, as shown in Figure 3. In contrast, the discrepancy between the four models was considered small for the accidents dataset as well as for the other datasets. When processing $50,000$ of traffic data from the accidents dataset, the difference in training runtime between the SISIV and baseline algorithms exceeds 300 seconds. These results may be explained by the fact that the baseline solutions use methods that mix deep learning architectures for feature extraction and basic machine learning algorithms for prediction, as well as ineffective hyperparameter tuning strategies that do not lead to optimal results. In contrast, SISIV's branch-and-bound algorithm with the combination of GCN and attention mechanism efficiently selects the parameters of the model as well as the relevant features of the input data, which reduces the training time.

TABLE II
F-MEASURE, PRECISION, RECALL PERFORMANCES.

| Dataset | Methods | P (%) | R (%) | F (%) |
|---|---|---|---|---|
| Astyx | RNN-LF | 53 | 37 | 44 |
| | TripRes | 51 | 64 | 57 |
| | GCN-LSTM | 52 | 60 | 56 |
| | Gra-TF | 71 | 66 | 68 |
| | IGCNN-RCD | 85 | 92 | 86 |
| Deep Drive | RNN-LF | 51 | 33 | 41 |
| | TripRes | 55 | 68 | 61 |
| | GCN-LSTM | 55 | 63 | 59 |
| | Gra-TF | 70 | 67 | 68 |
| | IGCNN-RCD | 84 | 91 | 87 |
| Landmarks | RNN-LF | 50 | 42 | 46 |
| | TripRes | 59 | 74 | 66 |
| | GCN-LSTM | 61 | 65 | 63 |
| | Gra-TF | 77 | 73 | 75 |
| | IGCNN-RCD | 86 | 99 | 92 |
| Accidents | RNN-LF | 52 | 48 | 50 |
| | TripRes | 61 | 77 | 68 |
| | GCN-LSTM | 65 | 69 | 67 |
| | Gra-TF | 81 | 78 | 79 |
| | IGCNN-RCD | 89 | 95 | 92 |

*3) F-measure, Precision, Recall Performances:* We conduct experiments with well-known traffic forecasting benchmarks such as Astyx, Deep Drive, Landmarks, and Accidents to demonstrate the superiority of the proposed framework in terms of precision, recall, and F-measure. Four baseline models were selected for comparison (RNN-LF, TripRes, GCN-LSTM and Gra-TF). The numerical results are presented in Table II. This table shows that the proposed framework performs better than the baseline solutions in every case. These

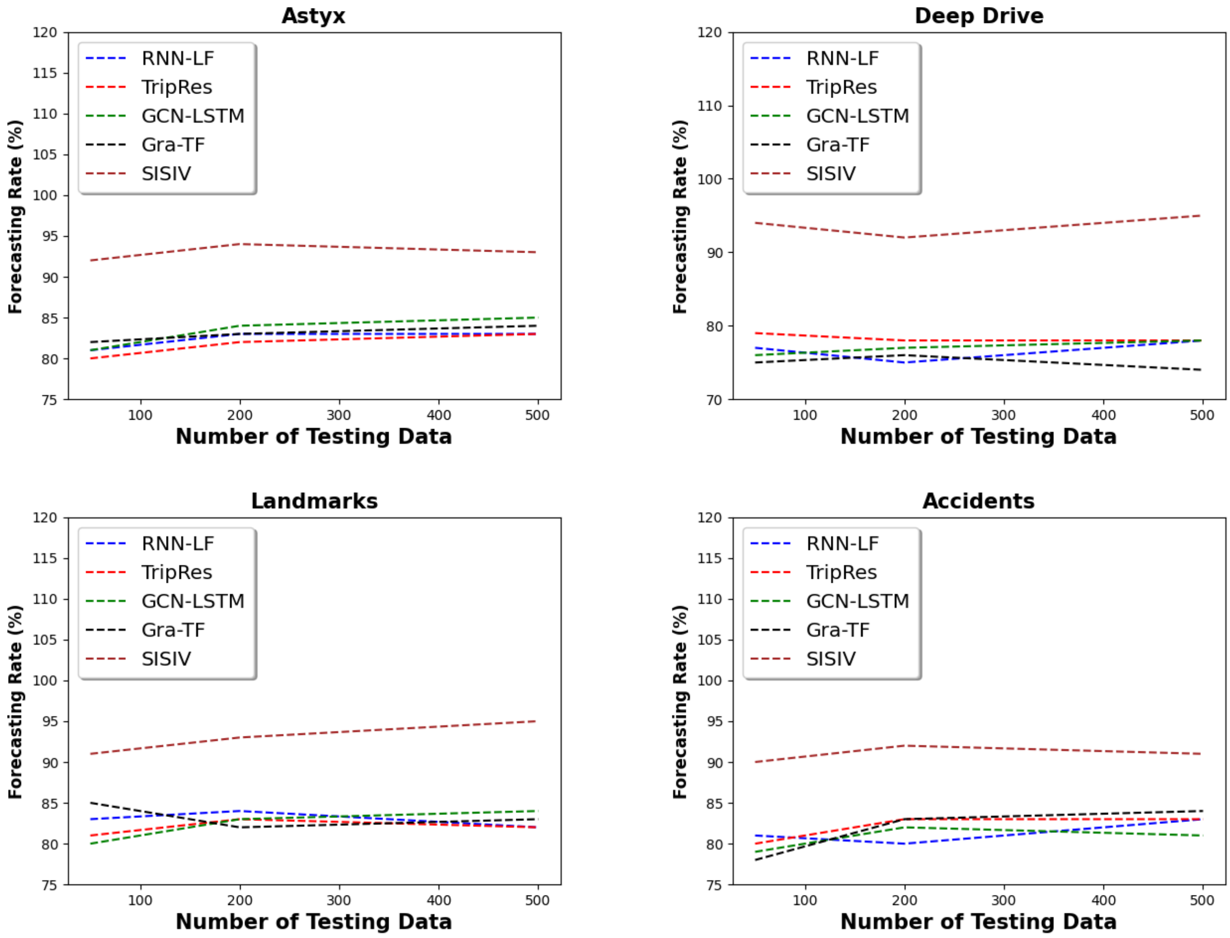


Fig. 2. The forcasting rate of SISIV compared to the baseline solutions.

results are the consequence of an effective mixture of graph optimization, GCN training and hyperparameter optimization. Graph optimization produces a densely packed graph of road networks. This enables efficient training of the GCN network. Moreover, the ideal model for traffic forecasting can be identified by exploring the GCN parameter space.

*4) Blockchain Performance:* The blockchain algorithm used in this study is evaluated in this part. Several tests were conducted using the traffic statistics presented in the previous section. Figure 4 shows the effect of different proportions of malicious vehicles on the successful attack rate, both with and without the blockchain technology used in this study. The results show that there is a demonstrable benefit of using the proposed blockchain method to secure various communications between vehicles in the transportation network.

## V. Open Research Scope for IoV

Research in IoV is growing by leaps and bounds. Compared to existing intelligent transportation technologies, these systems have recently been developed as IoV to remotely monitor vehicle operations and key parameters. It is only a matter of time before centralized IoV systems are improved and contribute significantly to reducing traffic accidents, saving maintenance costs, extending vehicle life, and increasing passenger and pedestrian safety. However, there are other problems and areas that need to be studied on the way to a reliable and efficient centralized IoV. Some of these studies are described below:

1) **A comprehensive investigation of vehicles:** A vehicle consists of thousands of parts. However, due to their individual and unique functions, not all of these parts are given the same importance. During the operation period, some of them take the main responsibility and ensure that the condition of the vehicle is satisfactory and also provide enhanced services. However, in our literature review, we found that previous research has focused exclusively on the functionality of critical vehicle elements as well as their impact on vehicle performance. There have been no mathematical advances in prioritizing these critical elements for use in this monitoring system. Therefore, there is great potential for research into a mathematical model for prioritizing vehicle components that could aid in the

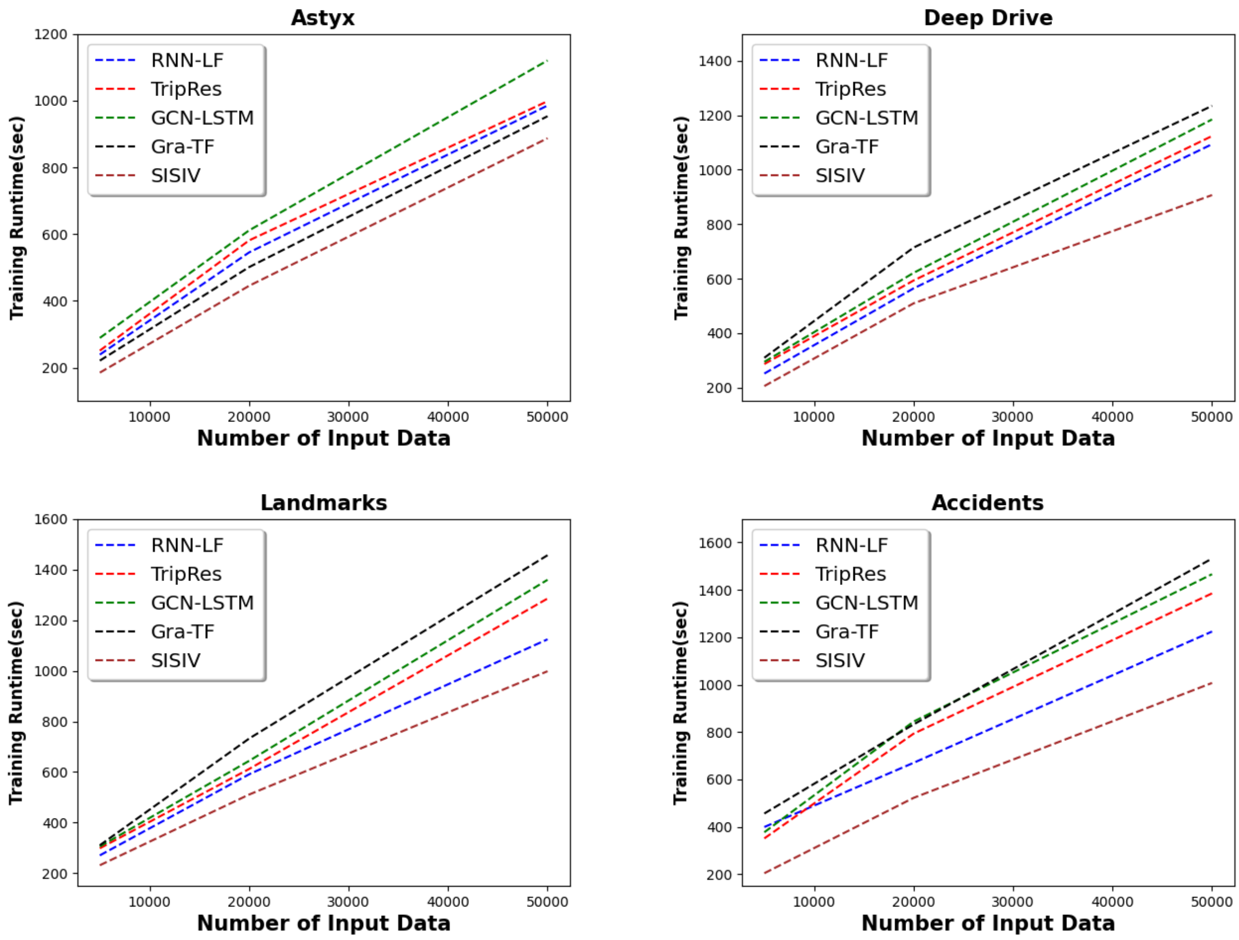


Fig. 3. SISIV's training runtime in comparison with the baseline solutions.

development of the IoV.

2) **Intelligent information fusion:** Apart from the security as well as service difficulties, the IoV system should be user-friendly and consider the preferences and privacy of all kinds of users. This is very important to arouse the interest of the users and keep them in supplying the systems with data. The research to develop the IoV core is crucial in this regard to make it easier for everyone by reducing the complexity of the system and enabling intelligent sharing and fusion of information.
3) **Optimized collaboration framework for IoV:** The number of connected vehicles is growing rapidly, and the ability to monitor them remotely is also growing in lockstep. Moreover, IoV is becoming increasingly important for remote monitoring of vehicle performance and operation. However, there is a problem with current networks, which have many limitations when it comes to connecting a large number of vehicles with roadside units, installation devices, surveillance systems, intelligent transportation systems, cloud storage and server systems, etc. To achieve this, IoV requires a well-organized and efficient communication network system that ensures a stable communication platform for vehicles to collect and analyze large amounts of data, as well as a platform for sharing data using IoT technology. A heterogeneous network is a viable alternative that can effectively meet the need. Developing an efficient and well-organized network for IoV that connects various links and nodes as a common platform holds great potential. In addition, combining exact and stochastic solutions could be a good direction to process large amounts of data in real-time operation.

## VI. Conclusion

This study examines the shortcomings of current IoV solutions as well as proposes the SISIV framework. An attention technique as well as a deep learning architecture based on GCNs are used. SISIV uses blockchain technology to secure the data exchange between nodes. Moreover, a branch-and-bound technique is used to intelligently determine the hyperparameters of the deep learning model created. The validation of SISIV was conducted using four networked vehicle databases designed to predict various traffic information. The results show that SISIV performs better

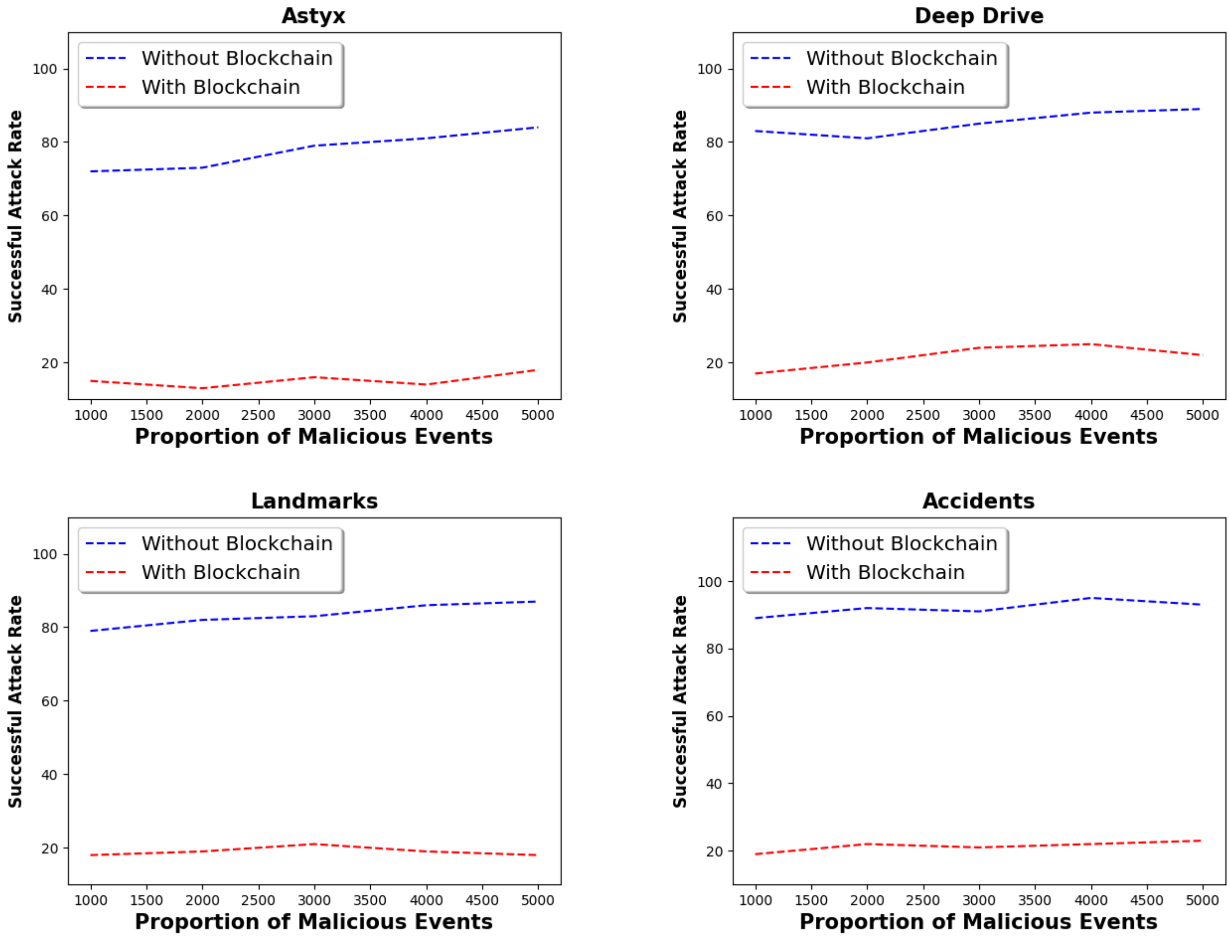


Fig. 4. Compare the influence of various percentage of malicious vehicles on the success rate of detected attacks with as well as without the use of Blockchain technology.

than the baseline solutions in terms of forecasting rate, F-measure, and detected attacks. There are many paths for future research productivity that emerge from the research conducted in this paper. First of all, IoV has made progress in many different functional areas, but security concerns are still a major problem for users of such systems [46]. For traffic forecasting, it is advisable to ensure a high level of security when data is retrieved and transmitted from users' devices. There has also been much recent research in the area of federated learning [47]. Applying the DL techniques explored in this paper in a FL based environment would be worthwhile considering the number of devices that would operate at the edge of such IoV networks.

## References


[1] F. Firouzi, B. Farahani, and A. Marinšek, "The convergence and interplay of edge, fog, and cloud in the ai-driven internet of things (iot)," *Information Systems*, vol. 107, p. 101840, 2022.

[2] E. Nehme, R. El Sibai, J. Bou Abdo, A. R. Taylor, and J. Demerjian, "Converged ai, iot, and blockchain technologies: a conceptual ethics framework," *AI and Ethics*, vol. 2, no. 1, pp. 129–143, 2022.

[3] S. Qazi, B. A. Khawaja, and Q. U. Farooq, "Iot-equipped and ai-enabled next generation smart agriculture: A critical review, current challenges and future trends," *IEEE Access*, 2022.

[4] A. Belhadi, Y. Djenouri, G. Srivastava, D. Djenouri, J. C.-W. Lin, and G. Fortino, "Deep learning for pedestrian collective behavior analysis in smart cities: A model of group trajectory outlier detection," *Information Fusion*, vol. 65, pp. 13–20, 2021.

[5] J. C.-W. Lin, G. Srivastava, Y. Zhang, Y. Djenouri, and M. Aloqaily, "Privacy-preserving multiobjective sanitization model in 6g iot environments," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5340–5349, 2020.

[6] T. Alladi, V. Kohli, V. Chamola, and F. R. Yu, "Securing the internet of vehicles: A deep learning-based classification framework," *IEEE networking letters*, vol. 3, no. 2, pp. 94–97, 2021.

[7] T. Alladi, V. Chamola, N. Sahu, V. Venkatesh, A. Goyal, and M. Guizani, "A comprehensive survey on the applications of blockchain for securing vehicular networks," *IEEE Communications Surveys & Tutorials*, 2022.

[8] D. Djenouri, R. Laidi, Y. Djenouri, and I. Balasingham, "Machine learning for smart building applications: Review and taxonomy," *ACM Computing Surveys (CSUR)*, vol. 52, no. 2, pp. 1–36, 2019.

[9] A. Belhadi, Y. Djenouri, G. Srivastava, D. Djenouri, A. Cano, and J. C.-W. Lin, "A two-phase anomaly detection model for secure intelligent transportation ride-hailing trajectories," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4496–4506, 2020.

[10] U. Ahmed, G. Srivastava, Y. Djenouri, and J. C.-W. Lin, "Deviation point curriculum learning for trajectory outlier detection in cooperative intelligent transport systems," *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[11] V. Hassija, V. Gupta, S. Garg, and V. Chamola, "Traffic jam probability estimation based on blockchain and deep neural networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp.

3919–3928, 2020.

[12] R. Wang, Y. Zhang, G. Fortino, Q. Guan, J. Liu, and J. Song, "Software escalation prediction based on deep learning in the cognitive internet of vehicles," *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS*, 2022.

[13] G. Srivastava, J. C.-W. Lin, A. Jolfaei, Y. Li, and Y. Djenouri, "Uncertain-driven analytics of sequence data in iocv environments," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 8, pp. 5403–5414, 2020.

[14] T. Alladi, V. Kohli, V. Chamola, F. R. Yu, and M. Guizani, "Artificial intelligence (ai)-empowered intrusion detection architecture for the internet of vehicles," *IEEE Wireless Communications*, vol. 28, no. 3, pp. 144–149, 2021.

[15] M. A. Rahman, M. A. Rahim, M. M. Rahman, N. Moustafa, I. Razzak, T. Ahmad, and M. N. Patwary, "A secure and intelligent framework for vehicle health monitoring exploiting big-data analytics," *IEEE Transactions on Intelligent Transportation Systems*, 2022.

[16] M. El Zorkany, A. Yasser, and A. I. Galal, "Vehicle to vehicle "v2v" communication: scope, importance, challenges, research directions and future," *The Open Transportation Journal*, vol. 14, no. 1, 2020.

[17] L. Xu, X. Zhou, Y. Fu, G. Jiang, X. Yu, M. Yu, N. Kumar, and M. Guizani, "Accurate and efficient performance prediction for mobile iov networks using gwo-gr neural network," *IEEE Internet of Things Journal*, 2022.

[18] B. Ji, Z. Chen, S. Mumtaz, C. Han, C. Li, H. Wen, and D. Wang, "A vision of iov in 5g hetnets: Architecture, key technologies, applications, challenges, and trends," *IEEE Network*, 2022.

[19] M. Zhang, J. Zhou, P. Cong, G. Zhang, C. Zhuo, and S. Hu, "Lias: A lightweight incentive authentication scheme for forensic services in iov," *IEEE Transactions on Automation Science and Engineering*, 2022.

[20] G. Wang, F. Xu, H. Zhang, and C. Zhao, "Joint resource management for mobility supported federated learning in internet of vehicles," *Future Generation Computer Systems*, vol. 129, pp. 199–211, 2022.

[21] S. Li, B. Wang, S. Qian, Y. Sun, X. Yun, and Y. Zhou, "Influence maximization for emergency information diffusion in social internet of vehicles," *IEEE Transactions on Vehicular Technology*, 2022.

[22] X. Liu, B. Lai, B. Lin, and V. C. Leung, "Joint communication and trajectory optimization for multi-uav enabled mobile internet of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, 2022.

[23] L. Nie, Y. Wu, X. Wang, L. Guo, G. Wang, X. Gao, and S. Li, "Intrusion detection for secure social internet of things based on collaborative edge computing: A generative adversarial network-based approach," *IEEE Transactions on Computational Social Systems*, 2021.

[24] T. Wang, P. Wang, S. Cai, X. Zheng, Y. Ma, W. Jia, and G. Wang, "Mobile edge-enabled trust evaluation for the internet of things," *Information Fusion*, vol. 75, pp. 90–100, 2021.

[25] S. M. Nagarajan, G. G. Deverajan, P. Chatterjee, W. Alnumay, and U. Ghosh, "Effective task scheduling algorithm with deep learning for internet of health things (ioht) in sustainable smart cities," *Sustainable Cities and Society*, vol. 71, p. 102945, 2021.

[26] A. Zekry, A. Sayed, M. Moussa, and M. Elhabiby, "Anomaly detection using iot sensor-assisted convlstm models for connected vehicles," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*. IEEE, 2021, pp. 1–6.

[27] M. Aloqaily, S. Otoum, I. Al Ridhawi, and Y. Jararweh, "An intrusion detection system for connected vehicles in smart cities," *Ad Hoc Networks*, vol. 90, p. 101842, 2019.

[28] Y. Fu, C. Li, F. Yu, T. H. Luan, and Y. Zhang, "A survey of driving safety with sensing, vehicular communications, and artificial intelligence-based collision avoidance," *IEEE Transactions on Intelligent Transportation Systems*, 2022.

[29] Z. Lu, G. Qu, and Z. Liu, "A survey on recent advances in vehicular network security, trust, and privacy," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 2, pp. 760–776, 2019.

[30] D. Djenouri, W. Soualhi, and E. Nekka, "Mobility models in vehicular ad hoc networks: the overtaking impact," *The Mediterranean Journal of Computers and Networks*, vol. 4, no. 2, pp. 62–70, 2008.

[31] S. Subramanian, D. Djenouri, G. Sindre, and I. Balasingham, "Cop4v : Context-based protocol for vehicle's safety in highways using wireless sensor networks," in *Sixth International Conference on Information Technology: New Generations, ITNG 2009, Las Vegas, Nevada, USA, 27-29 April 2009*. IEEE Computer Society, 2009, pp. 613–618.

[32] D. Djenouri, "Preventing vehicle crashes through a wireless vehicular sensor network," in *2008 24th IEEE Biennial Symposium on Communications, Kingston, Canada,*. IEEE Computer Society, 2008, pp. 320–323.

[33] A. A. Abdellatif, C. F. Chiasserini, F. Malandrino, A. Mohamed, and A. Erbad, "Active learning with noisy labelers for improving classification accuracy of connected vehicles," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 4, pp. 3059–3070, 2021.

[34] C. Li, S. Wang, X. Li, F. Zhao, and R. Yu, "Distributed perception and model inference with intelligent connected vehicles in smart cities," *Ad Hoc Networks*, vol. 103, p. 102152, 2020.

[35] Y. Xing, C. Lv, X. Mo, Z. Hu, C. Huang, and P. Hang, "Toward safe and smart mobility: Energy-aware deep learning for driving behavior analysis and prediction of connected vehicles," *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[36] A. Belhadi, Y. Djenouri, D. Djenouri, and J. C.-W. Lin, "A recurrent neural network for urban long-term traffic flow forecasting," 2020.

[37] X. Xu, Z. Fang, L. Qi, X. Zhang, Q. He, and X. Zhou, "Tripres: Traffic flow prediction driven resource reservation for multimedia iov with edge computing," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 17, no. 2, pp. 1–21, 2021.

[38] H. Peng, B. Du, M. Liu, M. Liu, S. Ji, S. Wang, X. Zhang, and L. He, "Dynamic graph convolutional network for long-term traffic flow prediction with reinforcement learning," *Information Sciences*, vol. 578, pp. 401–416, 2021.

[39] X. Xu, Z. Fang, J. Zhang, Q. He, D. Yu, L. Qi, and W. Dou, "Edge content caching with deep spatiotemporal residual network for iov in smart city," *ACM Transactions on Sensor Networks (TOSN)*, vol. 17, no. 3, pp. 1–33, 2021.

[40] D. Sun, J. Wu, J. Yang, and H. Wu, "Intelligent data collaboration in heterogeneous-device iot platforms," *ACM Transactions on Sensor Networks (TOSN)*, vol. 17, no. 3, pp. 1–17, 2021.

[41] K. Guo, Y. Hu, Z. Qian, H. Liu, K. Zhang, Y. Sun, J. Gao, and B. Yin, "Optimized graph convolution recurrent neural network for traffic prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 2, pp. 1138–1149, 2020.

[42] S. Luan, R. Ke, Z. Huang, and X. Ma, "Traffic congestion propagation inference using dynamic bayesian graph convolution network," *Transportation Research Part C: Emerging Technologies*, vol. 135, p. 103526, 2022.

[43] G. Chen, F. Lu, Z. Li, Y. Liu, J. Dong, J. Zhao, J. Yu, and A. Knoll, "Pole-curb fusion based robust and efficient autonomous vehicle localization system with branch-and-bound global optimization and local grid map method," *IEEE Transactions on Vehicular Technology*, 2021.

[44] 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.

[45] Q. Zhang, K. Yu, Z. Guo, S. Garg, J. Rodrigues, M. M. Hassan, and M. Guizani, "Graph neural networks-driven traffic forecasting for connected internet of vehicles," *IEEE Transactions on Network Science and Engineering*, pp. 1–1, 2021.

[46] M. Kamal, G. Srivastava, and M. Tariq, "Blockchain-based lightweight and secured v2v communication in the internet of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 3997–4004, 2020.

[47] V. Mothukuri, P. Khare, R. M. Parizi, S. Pouriyeh, A. Dehghantanha, and G. Srivastava, "Federated learning-based anomaly detection for iot security attacks," *IEEE Internet of Things Journal*, 2021.