

# Smart Internet of Vehicles Architecture based on Deep Learning for Occlusion Detection

Shaya A. Alshaya

Department of Computer Science

College of Science and Humanities at alGhat, Majmaah University, Majmaah, 11952, Saudi Arabia

**Abstract**—In our days, the cyber world is developing due to the revolution of smart cities and machine learning technologies. The internet of Things constitutes the essential background of cyber technology. As a case study, the Internet of Vehicles is one of the leading applications which is developed quickly. Studies are focused on resolving issues related to real-time problems and privacy leakage. Uploading data from the cloud during the data collection step is the origin of delay issues. This process decreases the level of privacy. The objective of the present paper is to ensure a high level of privacy and accelerated data collection. During this study, we propose an advanced Internet of Vehicle architecture to conduct the data collection step. An occlusion detection application based on a deep learning technique is performed to evaluate the IoV architecture. Training data at Distributed Intelligent layer ensures not only the privacy of data but also reduces the delay.

**Keywords**—Internet of vehicle; deep learning; collaborative technologies; cloud; edge computing

## I. INTRODUCTION

The recent development of network technology gives birth to cyber intelligence technologies. The Internet of Vehicles (IoV) is a subfield of the Internet of Things (IoT) that has evolved a new generation of network technology [1].

The IoV case analyzes data based on artificial intelligence (AI) and Machine Learning (ML) methods. This field is still challenged by analyzing big data, privacy, and computational power limitation.

The IoV cooperates with the cloud platform to ensure permanent services. Two issues are faced with the IoV applications: (1) delay problem and (2) privacy violation [2].

We define an intermediate layer called edge computing to manage the vehicle network layer and cloud layer to overcome these problems. The edge computing layer is determined by the location information and the low latency. This is obtained because edge nodes are constituted by the Road Side Unit (RSU) [3].

The edge computing layer indeed solves the above problems, but other difficulties are bred as (1) increase in the amount of data, (2) an increase in the amount of data types, (3) difficult to adapt existing data collection related to embedded data, (4) need for intelligent methods to collect data, and (5) need for smart methods for the transmission [4].

Deep learning methods provide a solution for the above difficulty. They ensure the simulation of the data analysis to

avoid forged data based on interpretation mechanisms. Besides, deep learning methods reduce redundancy and provide high robustness of the system by following the inspection's high quality.

Therefore, moving from a centralized approach to edge computing reduces the unnecessary jumps to the network and improves latency [5]. Privacy is enhanced by adopting a Collaborative Learning (CL) technology. It manages data through multiple decentralized edge devices without exchanging them. It follows a training method different from the traditional cloud data center. The trained data were not collected directly from terminals [6]. The CL technology ensures privacy by gathering models and their updates from each RSU.

All RSU shares the model of the cloud data center. The CL performs the following steps: (1) the user provides data, (2) the local model trains the provided data, and (3) the trained model is uploaded to the cloud data center.

However, the CL mechanism provides more privacy and reduces delay by decreasing the training time of the sharing model in the cloud.

In light of this introduction, we can summarize contributions into three points:

- 1) Propose an advanced architecture at the edge computing layer for the IoV. The proposed ensures intelligent data collection.
- 2) Propose a deep learning method for the preprocessing step. The proposed reduces the delay.
- 3) Deploy the CL technology between the local model from the edge node and the cloud edge.

The next section presents the literature review. The proposed IoV architecture is introduced in Section III. The data preprocessing schemes designed according to a deep learning method and collaborative learning are described in Section IV. Results and analyzes are discussed in Section V. A conclusion is highlighted at the end of this article.

## II. LITERATURE REVIEW

This section attempts to summarize the latest research related to the Internet of Vehicles based on the following directives: (1) The assessment of the quality of services, (2) occlusion detection.

### A. Quality of Service Assessment

The safety of the road and the crowd's reduction is enhanced according to the evaluation of the QoS [7]. Van der Lee et al., [8] propose a scheduling method to evaluate the vehicle network QoS. This method was named a Time Synchronized Channel Hopping (TSCH). It introduces the interference diagram, which is dependent on internal interference and conflict. This attempt provides accurate analyzes of the network performance. The ratio of packet reception and latency compose the performance metrics.

Zhang et al., [9] introduce an open access geometry-based efficient propagation model. This model is composed of a connected vehicle in the traffic network that analyzes the QoS according to the principal component, multidimensional scaling, and variance. The QoS is carrying out for vehicle-infrastructure and vehicle-vehicle.

### B. Occlusion Detection in IoV Systems

Avoid occlusion in traffic is the subject of many related works. Determine the density of the crowded in urban is still challenged [10]. The process aims to detect vehicles and exploit the crowded density. This is ensured through preprocessing, motion detection, feature extraction, and classification steps [11]. There are many traditional and new techniques used for classification [12]. Deep learning is the kernel of these new approaches like CNN approaches.

In the traditional approaches, features are determined manually, and the classification step computes the similarity according to these predetermined features. For example, the authors in [24] applied HOG algorithm to detect humans. Hadjkacem et al. [13] used Gait-Appearance-based Multi-Scale Video Covariance (GAMS-ViCov) to detect a pedestrian. The estimation of a high-dense crowd through images is the main subject discussed in [14]. The author used Balanced Communication-Avoiding Support Vector Machine classifier for the detection. In [15], the authors are focused on the case of the detection of vehicles by applying Harr-like and Adaboost features. The limitation of representation is the main issue of the traditional approaches.

In the deep learning approach, features are determined automatically according to the provided dataset. This approach offers a broad representation ability [16]. We focus on related works based on CNN with two-stage because it enriches higher performance compared to the one-stage approach [17][18]. The CNN-two-stage includes a region-based convolutional neural network [19]. Martinez et al., [19] perform CNN as a selective search. The purpose is to detect two thousand candidate regions through a fixed-sized input image. The obtained result in terms of performance is higher than the traditional approach but this method is time-consuming. The authors in [20] and [21] attempt to reduce the time consuming. They used the selective once (SPP-Net method) to remove it (Fast R-CNN method) as described in [21]. This led to announce the region network method instead of selective search. The region network ensures convolutional operation and sharing computation. These advancements provide a higher accuracy at minimum time-consuming.

## III. PROPOSED IOV ARCHITECTURE SCHEME

We introduce the proposed IoV architecture for the collection and transmission of data on an edge computing layer during this section. This architecture is composed of four sub-layers, as described in Fig. 1.

**Data collection layer:** This layer is built based on nodes. RSUs collect data related to vehicles and roads as vehicle location and traffic information [22]. These data are sent continuously to the distributed intelligence layer through edge devices.

**Distributed intelligence layer:** This layer aims to ensure preprocessing and data analysis received from the data collection layer. This intermediate layer between RSU and centralized cloud computing provides a powerful step to increase storage capabilities, share communication resources, and improve computing. The distributed intelligence layer is responsible for data transfer methods and the network's evaluation [23].

**Data processing layer:** In this layer, the collecting sensing data are trained. Three functions are verified in this step: (1) Detect the data quality, (2) Detect data similarity, and (3) Detect relevant data. This step decreases the amount of data transmitted to the cloud. Only the training results are sent to the cloud instead of direct transmission to the centralized cloud. The data processing layer reduces delays related to communications and increases the privacy level [24].

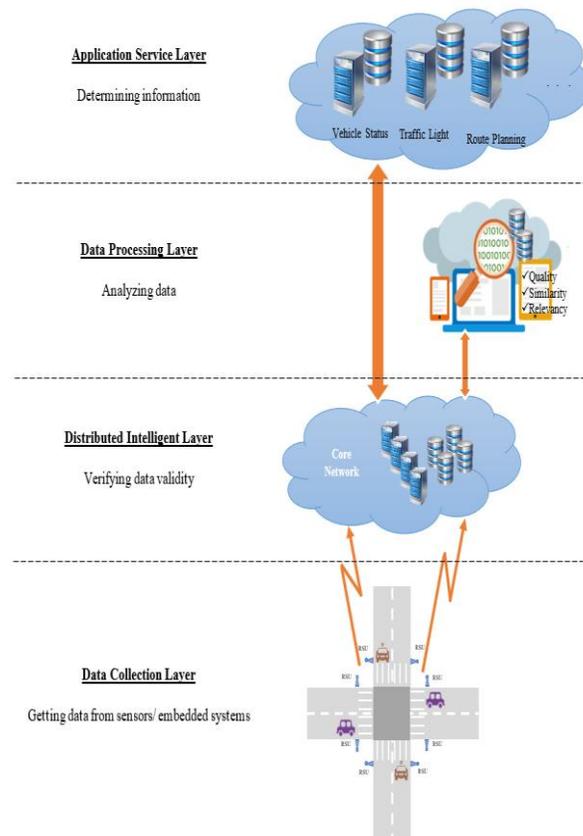


Fig. 1. Proposed IoV Architecture.

Application service layer: This layer offers information and helps make a decision based on the results obtained from sensing data. It monitors vehicle status, manages traffic lights, and updates route planning and directories.

Road sensors and systems embedded in vehicles transmit data to the RSU. The data collected by RSUs are moved to the core network. In this layer, data are conducted based on types. The data are verified according to the redundancy, quality, and relevancy at the data processing layer. When the verification is done, feedback is sent back to the distributed intelligent layer. Finally, the application service layer receives trained information. According to these results, the IoV systems are managed.

The proposed IoV architecture is built in terms of the network environment. The IoV model took into consideration network traffic and road environment. We carry out a strategy based on adaptive upload of the network bandwidth, environment, and transmission delay to enhance the collection data layer's performance.

The model divided the domain to ensure the Qos requirements. Divisions conditions are described by equation (1).

$$\begin{cases} D_i \in [x_i, y_i] \\ A_0 = \frac{x_0+y_0}{2} \\ A_i = \frac{x_i+y_i}{6} \\ H_e = Const \end{cases} \quad (1)$$

Therefore, building standards for the targeted network QoS evaluation was a necessity [25]. The Assessment standard is constructed by transforming qualitative data into quantitative data [26]. The network QoS is defined by the next parameters: namely, assessment indexes, bandwidth, delay, and rate of the lost packet.

Assessment indexes are composed of five levels of standard. Digital features are computed using the conditions presented in equation 1. Table I presents QoS features in the edge model.

TABLE I. DIGITAL FEATURES OF QOS ASSESSMENT IN EDGE MODEL

Assessment Level	Division	Digital features
Excellent	0.8-1	1.0,0.2 / 3,0.02
Good	0.65-1	0.78,0.15 / 5, 0.05
Average	0.5-0.65	0.64-0.1 / 5, 0.05
Poor	0.25-0.5	0.45-0.2 / 5,0.05
Very Poor	0.0.25	0,0.25 / 3,0.02

Then, we look for the optimal weight value. To achieve the target, we compute each secondary index's weight by applying the entropy method and the analytic hierarchy process. Then the square model is performed as described in equation 2.

$$\min g(w) = \sum_{i=1}^m \sum_{j=1}^n w [(s_j - w_j)p_{ij}]^2 + [(O_j - w_j)p_{ij}]^2 \quad (2)$$

Were  $s = [s_1 s_2 \dots s_n]^T$  is the computed weight through subjective analysis method,  $O = [O_1 O_2 \dots O_n]^T$  is the

computed weight through objective analysis method  $w = [w_1 w_2 \dots w_n]^T$  is the combined weight, the matrix  $P(p_{ij})^{m \times n}$  is the result of the standardized process of the measured data,  $m$  is the number of data, and  $n$  is the number of indicators.

The Lagrangian function is applied to reach the optimal weight, as mentioned in equations 3 and 4.

$$w = M^{-1} [B + \frac{1-E^T M^{-1} B}{I^T M^{-1} I} I] \quad (3)$$

Were

$$\begin{cases} M = diagonal [\sum_{i=1}^m p_{i1}^2, \sum_{i=1}^m p_{i2}^2, \dots, \sum_{i=1}^m p_{in}^2] \\ I = [1, 1, \dots, 1]^T \\ B = [\sum_{i=1}^m \frac{(s_1+v_1)p_{i1}^2}{2}, \sum_{i=1}^m \frac{(s_2+v_2)p_{i2}^2}{2}, \dots, \sum_{i=1}^m \frac{(s_n+v_n)p_{in}^2}{2}] \end{cases} \quad (4)$$

Then, digital features related to each level index are computed to determine the evaluation of the target layer edge using the entropy method (equation 5).

$$\begin{cases} E_i = \frac{E_{i1}E_{n1w1} + E_{i2}E_{n2w2} + \dots + E_{iq}E_{nqwq}}{E_n} \\ H_e = \frac{H_{e1}E_{n1w1} + H_{e2}E_{n2w2} + \dots + H_{eq}E_{nqwq}}{E_n} \\ E_n = E_{n1w1} + E_{n2w2} + \dots + E_{nqwq} \end{cases} \quad (5)$$

Were  $E_n$  measures the fuzziness and the probability of qualitative concept,  $E_i$  defines the qualitative concept in the number domain space,  $H_e$  is the doubt measure value of  $E_n$ .

The achieved digital features define precisely the adequate level of the edge model highlighted in Table I.

Table I applies many upload strategies based on selected network states. During this step, the scale of data is sent. The network congestion evaluation in this phase is composed of three levels: Good, Average, and Poor.

Results of the network status vary based on upload strategies. The network status is good when the RSU processes the image. The status moves to average when little videos are transmitted to the cloud. When the IoV uploads images and much video, the network status becomes poor.

#### IV. OCCLUSION DETECTION APPLICATION

The type of collected data is so crucial for the privacy and the amount of transmitted data. The proposed IoV architecture is evaluated via an occlusion detection application. Privacy is ensured by using a deep learning method to avoid the storage of the input data.

The detection of occlusion in urban traffic presents the main objective of each urban traffic surveillance video. This application is the most suitable to evaluate the proposed IoV architecture. The RSU applies the Convolutional Neural Network (CNN) for the input video to detect urban traffic status. The training is ensured at the distributed intelligent layer.

During this section, we focus on the detection of occlusion using the CNN method. The occlusion is considered based on global and local characteristics targets to detect precisely the vehicle position. Fig. 2 shows the occlusion detection framework.

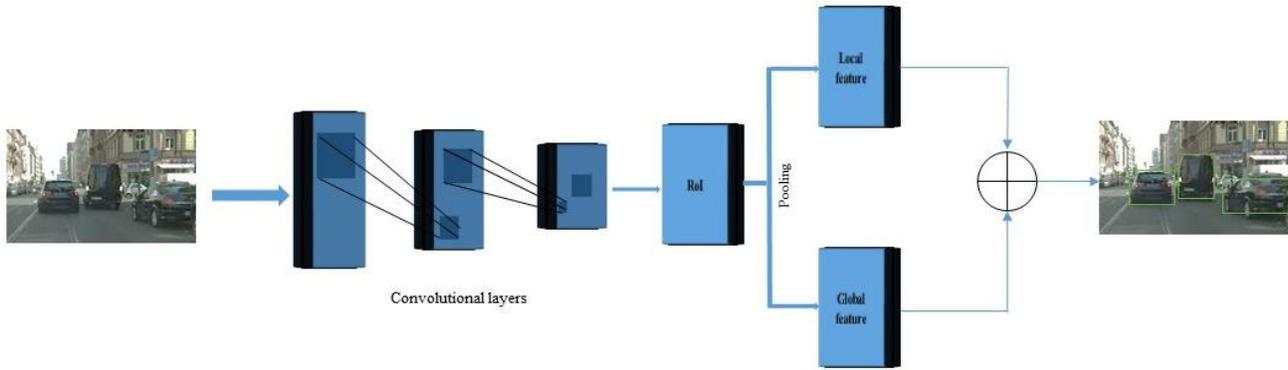


Fig. 2. Occlusion Detection Approach.

The local feature is extracted based on the position of the Region of Interest (RoI). A convolutional layer with size one and  $r^2 \times (C+1)$  channels are added to the last convolutional layer, where  $r$  defines the mask size, and  $C+1$  defines the background and the classes. The average of pooling related to the RoI indicates the probability of vehicle. Then, the bounding-box is labeled the same. Some channels are added to the bounding-box to compute the average pooling on the convolutional layer. The obtained result defines the values of bounding-box regression.

We consider the global feature of the video streaming of urban traffic enhances the accuracy of vehicle detection. The RoI pooling is used in this case to avoid errors related to the size of vehicles. Then, we apply  $1 \times 1$  convolutional layer to select the global feature.

The concatenation of these features provides accurate detection.

## V. EXPERIMENTATIONS

The proposed IoV architecture and the occlusion detection application are verified during this section. We start to display the achieved results related to the occlusion detection application.

The occlusion detection using a deep learning algorithm is tested based on UIUC dataset [27]. This dataset is composed of 550 grayscale image cars. Images are both single-scale and multi-scale. We use 400 images for training and the rest for testing.

The learning rate starts from 0.01, then changes to 0.001 after 18000 iterations. The training requires 302s to be performed. The training is characterized by momentum, which is about 0.92, and weight decay which is about 0.002.

In the complete evaluation of the proposed architecture, we use 150 images. We introduce two metrics: (1) True Positive Rate (TPR), and (2) False Positive Rate (FPR).

Fig. 3 shows the Receiver Operating Characteristics graph (ROC) of occlusion detection. The ROC is defined as the relationship between TPR and FPR. The ROC curve overgrows.

The best baseline highlighted in Fig. 3 is 0.8. In comparison with Shivani et al., in [27], we demonstrate that our method had the best baseline.

We drew also, the accuracy rate curve according to the baseline, as shown in Fig. 4. The curve is composed of one peak which is achieved at value 0.91. Then, the curve go-down and stabilizes under the accuracy rate value is 0.2. The baseline is selected at 0.9.

The processing requests 54.2 ms to be performed. The proposed architecture of data collection and transmission is evaluated according to the detection of occlusion in traffic.

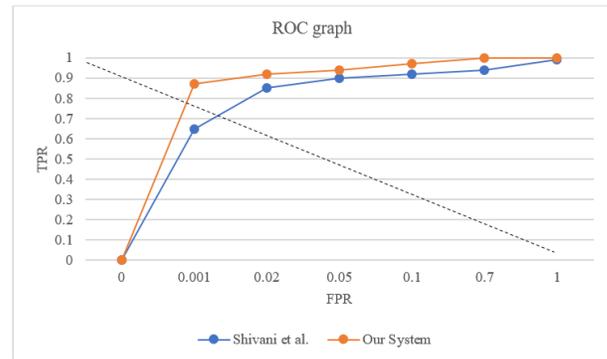


Fig. 3. ROC Curves.

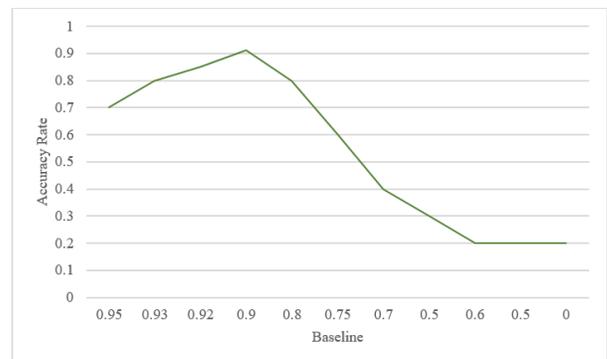


Fig. 4. Baseline and Accuracy of Image Traffic.

The recognition accuracy and latency is computed to study the performance in-depth. In our case, the evaluation is done through 550 images in which the number of cars per image varies between two and ten. The latency is computed using 2.4 G network environment. Its curve is reduced and did not exceed 6s, as shown in Fig. 5. In contrariwise, the recognition accuracy curve (Fig. 6) depends directly on the data size. The more the data bigger, is more the accuracy is improved. This is well justified due to the use of deep learning. The accuracy achieves a rare about 95%.

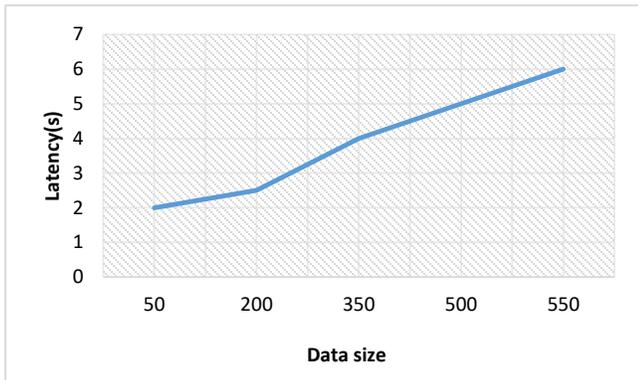


Fig. 5. Latency Curve.

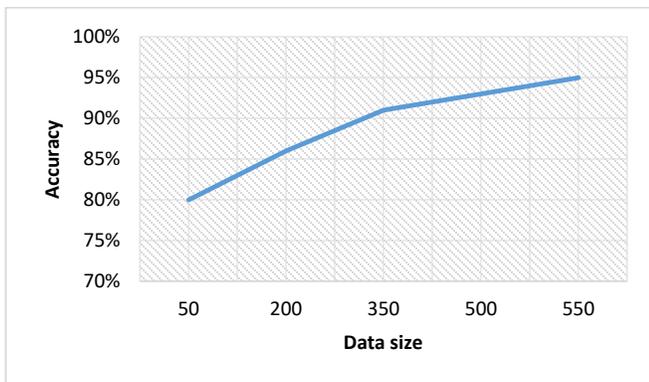


Fig. 6. Accuracy Curve.

The experimental results prove that the proposed IoV architecture obtains high performance in the detection of occlusions in decreasing delay and ensuring a high level of privacy.

## VI. CONCLUSION

The network intelligence technology is faced with the leakage of privacy problem. Provide a new preprocessing technique to detect redundant data is so important. Therefore, this paper uses deep learning and edge computation to design an accurate data collection and preprocessing scheme. The cloud data centers and edge devices are managed through collaborative learning technologies and deep learning models. The IoV architecture describes the enhancement of the adaptability and efficiency of data collection. Detection of occlusion is performed to verify the correlation of data. Achieved results highlight the reduction of data uploaded to the cloud and the safety of the user's privacy.

## ACKNOWLEDGMENT

The authors would like to thank Deanship of Scientific Research at Majmaah University for funding this project under the number R-2021-61.

## REFERENCES

- [1] T. Wang, H. Luo, X. Zheng, and M. Xie, "Crowdsourcing mechanism for trust evaluation in CPCS based on intelligent mobile edge computing," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 6, pp. 1–19, 2019.
- [2] Y. Wu, H. Huang, Q. Wu, A. Liu, and T. Wang, "A risk defense method based on microscopic state prediction with partial information observations in social networks," *J. Parallel Distrib. Comput.*, vol. 131, pp. 189–199, 2019.
- [3] X. Liu, T. Wang, W. Jia, A. Liu, and K. Chi, "Quick convex hull-based rendezvous planning for delay-harsh mobile data gathering in disjoint sensor networks," *IEEE Trans. Syst. Man, Cybern. Syst.*, 2019.
- [4] T. Wang, Y. Mei, W. Jia, X. Zheng, G. Wang, and M. Xie, "Edge-based differential privacy computing for sensor–cloud systems," *J. Parallel Distrib. Comput.*, vol. 136, pp. 75–85, 2020.
- [5] B. Yin, Y. Wu, T. Hu, J. Dong, and Z. Jiang, "An Efficient Collaboration and Incentive Mechanism for Internet of Vehicles (IoV) With Secured Information Exchange Based on Blockchains," *IEEE Internet Things J.*, vol. 7, no. 3, pp. 1582–1593, 2019.
- [6] K. Yan, W. Shen, Q. Jin, and H. Lu, "Emerging privacy issues and solutions in cyber-enabled sharing services: From multiple perspectives," *IEEE Access*, vol. 7, pp. 26031–26059, 2019.
- [7] T. Wang, H. Luo, W. Jia, A. Liu, and M. Xie, "MTES: An intelligent trust evaluation scheme in sensor-cloud-enabled industrial Internet of Things," *IEEE Trans. Ind. Informatics*, vol. 16, no. 3, pp. 2054–2062, 2019.
- [8] T. van der Lee, A. Liotta, and G. Exarchakos, "Time-scheduled network evaluation based on interference," in *2018 IEEE International Conference on Cloud Engineering (IC2E)*, 2018, pp. 323–332.
- [9] L. Zhang, Y. Liu, Z. Wang, J. Guo, and Y. Huo, "Mobility and QoS oriented 802.11 p MAC scheme for vehicle-to-infrastructure communications," *Telecommun. Syst.*, vol. 60, no. 1, pp. 107–117, 2015.
- [10] S. A. Alshaya, "Open Challenges for Crowd Density Estimation," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 1, 2020, doi: 10.14569/IJACSA.2020.0110123.
- [11] A. A. Yilmaz, M. S. Guzel, I. Askerbeyli, and E. Bostanci, "A vehicle detection approach using deep learning methodologies," *arXiv Prepr. arXiv1804.00429*, 2018.
- [12] S. Zhang, L. Wen, X. Bian, Z. Lei, and S. Z. Li, "Single-shot refinement neural network for object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4203–4212.
- [13] B. Hadjkacem, W. Ayedi, M. Ben Ayed, S. A. Alshaya, and M. Abid, "A novel Gait-Appearance-based Multi-Scale Video Covariance Approach for pedestrian (re)-identification," *Eng. Appl. Artif. Intell.*, vol. 91, p. 103566, 2020.
- [14] S. A. Alshaya, "Estimation of a high-dense crowd based on a Balanced Communication-Avoiding Support Vector Machine classifier," *Int. J. Comput. Sci. Netw. Secur.*, vol. 20, no. 6, pp. 195–201, Jun. 2020.
- [15] A. Tannouche, K. Sbai, M. Rahmoune, R. Agounoune, A. Rahmani, and A. Rahmani, "Real Time Weed Detection using a Boosted Cascade of Simple Features.," *Int. J. Electr. Comput. Eng.*, vol. 6, no. 6, 2016.
- [16] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, 2006.
- [17] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes challenge: A retrospective," *Int. J. Comput. Vis.*, vol. 111, no. 1, pp. 98–136, 2015.
- [18] T.-Y. Lin et al., "Microsoft coco: Common objects in context," in *European conference on computer vision*, 2014, pp. 740–755.

- [19] A. Alonso, M. Del Valle, J. A. Cecchini, and M. Izquierdo, "Asociación de la condición física saludable y los indicadores del estado de salud (II)," 2003.
- [20] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1904–1916, 2015.
- [21] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, 2015, pp. 91–99.
- [22] T. Wang, L. Qiu, A. K. Sangaiah, G. Xu, and A. Liu, "Energy-efficient and trustworthy data collection protocol based on mobile fog computing in Internet of Things," *IEEE Trans. Ind. Informatics*, vol. 16, no. 5, pp. 3531–3539, 2019.
- [23] Z. Xia, Z. Hu, and J. Luo, "UPTP vehicle trajectory prediction based on user preference under complexity environment," *Wirel. Pers. Commun.*, vol. 97, no. 3, pp. 4651–4665, 2017.
- [24] Q. Tang, M. Xie, K. Yang, Y. Luo, D. Zhou, and Y. Song, "A decision function based smart charging and discharging strategy for electric vehicle in smart grid," *Mob. Networks Appl.*, vol. 24, no. 5, pp. 1722–1731, 2019.
- [25] T. Wang, M. Z. A. Bhuiyan, G. Wang, L. Qi, J. Wu, and T. Hayajneh, "Preserving balance between privacy and data integrity in edge-assisted Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 4, pp. 2679–2689, 2019.
- [26] B. Xiong, K. Yang, J. Zhao, and K. Li, "Robust dynamic network traffic partitioning against malicious attacks," *J. Netw. Comput. Appl.*, vol. 87, pp. 20–31, 2017.
- [27] S. Agarwal, A. Awan, and D. Roth, "Learning to detect objects in images via a sparse, part-based representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 11, pp. 1475–1490, 2004.

#### AUTHOR'S PROFILE

Dr. Shaya Alshaya: Coming from the invaluable educational background in both under and postgraduate degrees from Saudi Arabia, USA, and Italy, Dr. Alshaya has put his knowledge into practicality as an international entrepreneur, his endeavors include education and research, ICT sector. He is a member of the National Association of Industrial Technology, the Institute of Electrical Engineers and the Sloan-C Online Association amongst others. Dr. Alshaya has a wealthy experience as Head of IT in Majmaah University, Saudi Arabia. His experience as a researcher, consultant, project manager and information technologist has allowed him to bring his extensive experience of technology planning, training and implementation to all of the projects he endeavors. The numerous national and international companies that utilize his expertise as a consultant evidence Dr. Alshaya's success. Dr. Alshaya's wealthy experience, vast knowledge is integral to his vision and future projects.