IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 13 May 2023, accepted 11 June 2023, date of publication 19 June 2023, date of current version 3 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3287488

RESEARCH ARTICLE

Colliding Bodies Optimization With Deep Belief Network Based Robust Pedestrian Detection

MAHA FAROUK SABIR[1], HADI OQAIBI[1], SAMI SAEED BINYAMIN[2], TURKI ALTHAQAFI[3], AND MAHMOUD RAGAB[4,5]

[1]Information Systems Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia
[2]Computer and Information Technology Department, The Applied College, King Abdulaziz University, Jeddah 21589, Saudi Arabia
[3]Information Systems Department, HECI School, Dar Al-Hekma University, Jeddah 22246, Saudi Arabia
[4]Information Technology Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia
[5]Mathematics Department, Faculty of Science, Al-Azhar University, Naser City, Cairo 11884, Egypt

Corresponding author: Mahmoud Ragab (mragab@kau.edu.sa)

This work was supported by the Deanship of Scientific Research (DSR) at King Abdulaziz University (KAU), Jeddah, Saudi Arabia, under grant no. (G: 465-612-1443).

ABSTRACT Pedestrian detection is a significant research topic in the computer vision (CV) domain for a longer period. Recently, deep learning (DL) and specifically convolutional neural network (CNN) exhibit significant improvement in the computer vision tasks such as object detection, segmentation, image classification, etc. With this motivation, this study develops a novel Colliding Bodies Optimization with Deep Learning based Robust Pedestrian Detection (CBODL-RPD) model. The goal of the CBODL-RPD approach is to identify the occurrence of pedestrians and non-pedestrians via object detection process. For object detection process, YOLO v4 with Adagrad optimizer is applied. In addition, the CBODL-RPD technique employs SqueezeNet model to generate feature vectors, and the hyperparameter tuning process is performed via the CBO algorithm. At last, deep belief network (DBN) model is applied for accurate pedestrian detection. A comprehensive experimental analysis is made to demonstrate the significant pedestrian detection results of the CBODL-RPD technique. The comparative outcome study reported the improved outcomes of the CBODL-RPD method over other recent methods.

INDEX TERMS Pedestrian detection, video analysis, computational intelligence, deep learning, parameter optimization.

I. INTRODUCTION

Recently, many videos and images are acquired through image acquisition gadgets or sensor measurements [1]. Such videos and images are utilized for tracking, detecting, and finding targets of interest, and the discoveries are employed for ecological conservation, plant monitoring, face detection, etc. Various effectual computing methods are emerged with the advances in computational intelligence which includes deep learning (DL) [2], expert systems, machine learning (ML) and other intellectual methods. Such methods are employed for video and image analysis, including decision making, feature extraction, or data mining [3]. But proper analyses and interpretations of the acquired information are difficult. Appropriate implementation or model of computational intelligence techniques will help enhance the speed and efficiency of video and image processing [4]. Artificial intelligence (AI), along with computer vision (CV) methods, intends to automatic interpretation of the visual context of scenes in the form of sequence of frames or a single frame and react consequently [5]. The difficulty that CV researchers have been facing until now is the recognition of pedestrians or human shapes. It becomes a significant element in several higher-level applications, from car security to advanced surveillance mechanisms [6]. Fostered by the arrival of highly robust yet compact hardware, the past few years have seen substantial advancements in pedestrian recognition systems in respect of effectiveness and accurateness [7].

Many pedestrian detection approaches share same computational pipelines [8]. Initially, commencing from raw pixel

The associate editor coordinating the review of this manuscript and approving it for publication was Aasia Khanum.

level image content, they would derive high level features or spatial representations resorting to randomly complicated conversion that can be implemented window-by-window or pixel-by-pixel. Secondly, features for any of the presented spatial windows were given to classifiers that evaluate whether that area would depict humans. Additionally, a scale space can be commonly employed for identifying pedestrians at distinct scales, i.e., distance relating to the sensing devices. Deep neural networks (DNNs) are rapidly transforming the world of AI and ML [9]. They set novel benchmarks for several heterogeneous applications in several fields, comprising natural language processing (NLP), image understanding, and speech and audio analysis, bridging the gap in human performance for various errands. To be specific, CNN) signified a revolution in image analysis [10]. It considers existing tasks such as object detection, image classification, and face recognition. For pedestrian detection, there was increasing interest in CNN for the past years, inspired by the same image analysis task achievements.

This study develops a novel Colliding Bodies Optimization with Deep Learning based Robust Pedestrian Detection (CBODL-RPD) model. The main aim of the CBODL-RPD algorithm is to perform pedestrian detection via object detection and hyperparameter tuning process. To do so, the proposed model uses YOLO v4 for object detection process with SqueezeNet model as hyperparameter optimizer. Besides, the hyperparameter tuning process is performed via the CBO algorithm. Finally, deep belief network (DBN) model is applied for accurate pedestrian detection. A brief set of simulations were carried out to investigate the enhanced pedestrian detection results of the CBODL-RPD technique. In summary, the key contributions of the paper is given as follows.

- A novel CBODL-RPD technique is developed for pedestrian detection by comprising YOLOv4 based object detector, SqueezeNet feature extraction, CBO based parameter tuning, and DBN classification.
- The parameter optimization of the SqueezeNet model using the CBO algorithm using cross- validation helps to boost the predictive outcome of the proposed model for unseen data.

The rest of the paper is organized as follows. Section II provides the related works and section III offers the proposed model. Then, section IV gives the result analysis and section V concludes the paper.

II. LITERATURE REVIEW

Song et al. [11] introduced a strong multispectral feature fusion network (MSFFN) for detecting pedestrians that completely incorporates the extracted features from infrared and visible light channels. The especially multiscale semantic feature was extracted through two key components, such as multiscale feature extraction of infrared images (MFEI) and multiscale feature extraction of visible images (MFEV), and combined with the amended YOLOv3 networks for recognizing pedestrians. In [12], selected a low power entrenched Graphics Processing Unit (Jetson Nano) that enables various neural networks (NN) to be parallelly carried out and a CV system to be employed for recognizing the image. Additionally, the performance of DL-NNs, namely pednet, ssd-inception v2, multiped, and ssd-mobilenet v1 and v2, was tested. Furthermore, it has shown that the processing time and accuracy are enhanced while every model recommended in the study was employed.

Hung et al. [13] developed a DL-based mechanism for pedestrian detection problems from a drone-oriented image. Especially the study used a Fast RCNN to fine the extant pedestrian inside drone-oriented image. To evaluate the performance, an overall of 1500 images were gathered using S30W drones, and such images were captured at distinct locations. Jiang et al. [14] introduced an ML classification that is utilized for pedestrian recognition based on XGBoost. Genetic algorithm (GA) is presented for enhancing the parameter tuning method in training the XGBoost architecture. To enhance the classifier performance, HOG and LBP features are utilized for describing pedestrians, later inputted into GA-XGBoost classification presented to construct a novel static image pedestrian recognition method. Kim et al. [15] designed a pedestrian recognition algorithm related to deep CNN for pedestrian classification from the input image. The study presents an enhanced version of the VGG-16 structure was assessed for identifying pedestrians on the INRIA benchmark dataset comprising 227×227-pixel images.

Liu et al. [16] proposed the deeply separable fusion hierarchical feature module into backbone network of the Single Shot MultiBox Detector (SSD) method that decreases complexity of the method through depth-wise Separable Convolution and employs hierarchical structure fusion to improve features. Simultaneously, the encoder was utilized to transmit the semantic feature and output shallow feature for pedestrian recognition, enhancing classification and recognition capability. Yang et al. [17] introduced a Part-Aware Multi-Scale Fully Convolutional Networks (PAMS-FCNs) for managing such problems. To be Specific, for mining body parts with distinct responses, the author presented a part-aware RoI pooling module and chose the part with powerful response by voting. Essentially, a partially visible pedestrian case could get maximum recognition confidence scores, making it less possible to be a missing recognition. This element will operate parallel with a RoI pooling component for integrating global context information and local parts.

In [18], the authors developed an intelligent multimodal pedestrian detection approach which involves YOLO based object detection and kernel extreme learning machine (KELM) algorithm for pedestrian classification. Finally, hybrid salp swarm optimization (HSSO) model is used for optimal parameter adjustment. The authors in [19] presented a new integration approach to estimate real time poses and track multiple humans in the pedestrian environment, specifically for social distances using surveillance videos. The

proposed model offers a steady rise in processing speed and enhanced detection in low-resolution scenario. The authors in [20] focused on the development of pedestrian detection model by the use of multiple cameras. DL based YOLO model is used for object detection and multi-object tracking algorithm (DeepSORT) can be utilized for pedestrian tracking process. In [21], a new lightweight model with improved performance is derived depending upon the YOLO-ECA and SiamCSP. In the method, the Efficient Channel Attention block (ECA) was introduced into the feature pyramid network (FPN) of YOLO to allocate resources more adaptively. The authors in [22] introduced an effective crowded pedestrian detection model named YOLO-CPD that operates effective over other one-stage methods in the crowded environment. The presented model boosts the capability of the one-stage detector for the detection of many overlapping objects in a single area.

Though numerous object detectors are available in the literature, there is a still a need to develop automated and accurate pedestrian detection models. The DL models gets deepened with more number of layers and can leads to model overfitting. To resolve this issue, hyperparameter tuning process is recommended. Particularly, the hyperparameters such as epoch count, batch size, and learning rate selection are essential to attain effectual outcome. Since the trial and error method for hyperparameter tuning is a tedious and erroneous process, metaheuristic algorithms can be applied. Therefore, in this work, we employ CBO algorithm for the parameter selection of the SqueezeNet model.

III. THE PROPOSED MODEL

In this study, we have developed a new CBODL-RPD algorithm for an accurate and efficient pedestrian detection process. The main goal of the CBODL-RPD method is to identify the occurrence of pedestrians and non-pedestrians via object detection process. It involves a series of processes (as shown in Fig. 1), namely YOLO-v4 object detection, Adagrad optimizer, SqueezeNet feature extraction, CBO hyperparameter tuning, and DBN classification.

A. OBJECT DETECTION MODULE: YOLO-v4

For object detection process, YOLO v4 with Adagrad optimizer is applied. YOLOv4 is the newest model of YOLO that is foundation of YOLOv3, scales up and down and was suitable for large and small networks whereas preserving optimum accuracy and speed [23]. In comparison to YOLOv3, YOLOv4-tiny was an expanded edition of YOLOv3. The original Darknet53 will be added to a CSP network. Backbone was CSPO- SANet developed by Cross Stage Partial Network (CSPNet+One-Shot Aggregation Network (OSANet), including Partial in Computation Block (PCB) technique. CSPNet is employed for distinct CNN structures to decrease the amount of calculations and parameters whereas enhancing the precision. OSANet was drawn from OSA mechanism in VoVNet.

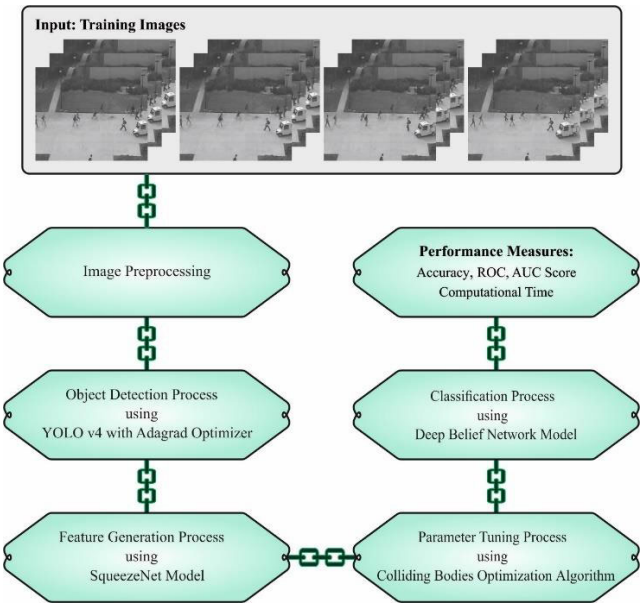

FIGURE 1. Working principle of CBODL-RPD model.

The key concept is enhanced using DenseNet architecture. Latterly, each layer is interconnected to enable input consistent with amount of output channels; PCB technique makes the method flexible since it is attuned based on the architecture to accomplish better accuracy-speed balance. The loss function rests the same as the YOLOv4, comprising confidence loss, regression loss, and classification loss [23]. Confidence loss and Classification loss remain same as the YOLOv3, however complete intersection over union (CIoU) was applied for replacing MSE to improve the regression loss:

$$\begin{aligned} LOSS = {} & 1 - IoU + \frac{\rho^2\left(b, b^{gt}\right)}{c^2} + \alpha v \\ & - \sum_{i=0}^{S2} \sum_{j=0}^{B} I_{ij}^{obj} \\ & \times [\hat{C}\log(C_i) + (1-\hat{C})log(1-c_i)] \\ & - \lambda_{noobj} \sum_{i=0}^{S2} \sum_{j=0}^{B} I_{ij}^{noobj} \\ & \times [\hat{C}_i log(C_i) + (1-\hat{C}_i)log(1-c_i)] \\ & - \sum_{i=0}^{S2} I_{ij}^{obj} \sum_{c \in classes} \\ & \times [\hat{p}_i(c)log(p_i(c)) + (1-\hat{p}_i(c))log(1-p_i(c))], \end{aligned} \tag{1}$$

From the expression, S^2 denotes $S \times S$ grids; every grid produces B candidate box, and every candidate box catches corresponding bounding box via the network; lastly, $S \times S \times B$ bounding box was created. The confidence loss of the box is evaluated if there was no object (*noobj*) in the box. The confidence loss function employed cross entropy error and was splitted into two parts: there exists the object (*obj*) and *noobj*. The loss of *noobj* rises weight coefficient λ that decreases contribution weight of the *noobj*. Also, classification loss function makes use of cross entropy error. Once the $j-th$

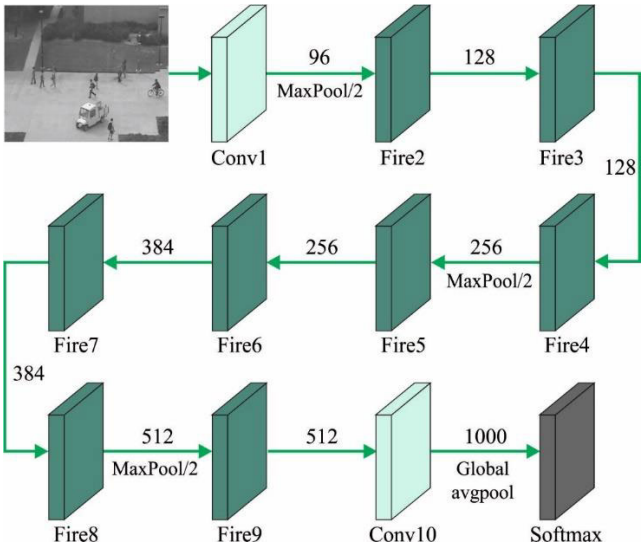

FIGURE 2. Structure of SqueezeNet.

anchor box of the $i-th$ grid is accountable for specific ground truth.

To adjust the YOLO v-4 hyperparameters, the Adagrad optimizer is applied. The Adagrad optimizer is utilized for hyperparameter tuning of the Mask RCNN mechanism. Evaluate gradient and accumulate squared gradient of every parameter for Adagrad as follows [24]:

$$G_t = \sum_{\tau=1}^{t} g_\tau \odot g_\tau \tag{2}$$

In Eq. (2), $\odot$ demonstrating a component-wise multiplication and $g_\tau \in \mathrm{R}^{|\theta|}$ represents the gradient of present variable at τ iteration. The upgraded value of variable in Adagrad is given by:

$$\Delta\theta_t = -\frac{\alpha}{\sqrt{G_t + \varepsilon}} \odot g_\tau \tag{3}$$

From the expression, α characterizes the rate of learning and ε signifies a smoothing term that avoids division by zero. Meanwhile, the learning rate has been determined beforehand training

$$\theta_t = -\alpha\left(\frac{1}{\sqrt{G_t + \varepsilon}} \odot g_\tau\right) \tag{4}$$

Now, G_t signify the prior gradient, accordingly, the expression enclosed in parentheses is regarded as one type of gradient revision

$$g_t' = \frac{1}{\sqrt{G_t + \varepsilon}} \odot g_\tau \tag{5}$$

Hence, the upgraded value of Adagrad is formulated as follows:

$$\Delta\theta_t = -\alpha g_t' \tag{6}$$

Traditional gradient descent is like an upgrading method. Therefore, Adagrad is regarded as an optimization method based on gradient.

B. FEATURE EXTRACTION MODULE

The CBODL-RPD technique employs SqueezeNet model to generate feature vectors. The SqueezeNet method is given in Fig. 2. The SqueezeNet technique [25] has 8 Fire modules, convolution layer, and final convolution layer. The SqueezeNet will use minimal variables which fit with the memory and makes process to operate effortlessly. It decreases the number of poses and attributes the capability to manage big data. Additionally, it will optimize the network and was computationally cost effective. In this study, feature vector Z was used as an input. The fire module has squeeze convolution layer that includes 1×1 filters that are imperilled to an expanded layer and has 3×3 and 1×1 convolution filters. The $\nu 1 \times 1$, $\chi 1 \times 1$ and $\chi 3 \times 3$ are the 3 tunable sizes in fire module subsist. The filters presented in the squeeze layers were designated as $\nu 1 \times 1$, $\chi 3 \times 3$ signifies the number of 3×3 filter in expand layer, and $\chi 1 \times 1$ denotes number of 1×1 filter in expanding layer.

Here, the hyperparameter tuning process is performed via the CBO algorithm. CBO is a metaheuristic approach inspired by the theory of $1D$ collision [26]. Each searching agent can be formed as a body having velocity and mass. The first position of $i - th$ body can be provided arbitrarily in the j-th parameter as follows:

$$x_{ij} = x_{j,\min} + rand \cdot \left(x_{j,\max} - x_{j,\min}\right), \tag{7}$$

In Eq. (7), rand signifies an arbitrary number ranging from zero to one. A collision increases amongst the location, and 2 bodies, afterwards impact, are upgraded along with 1D collision theory. Assume the body X_k (particle or object), its mass can be described as in Eq. (8):

$$m_k = \frac{1/J_k}{1/\sum_{i=1}^{n}(1/J_i)}, k = 1, \ldots, n \tag{8}$$

From the expression, J_k indicates cost function of $k - th$ particles and $n - th$,that must be even numbers, was the total volume of bodies exploited in optimizer technique. The n colliding body (CB) was sorted in ascending order, according to the values of objective function, and split up into Stationary Object (the lower half) and Moving Object (the upper half). The MO objects collide in contrast to member of SO groups to optimize the place and push stationary objects to the best place. Principally, the colliding pairs are defined according to ascending order. Hence, the optimum moving particle collide with optimum stationary one as follows:

$$Stationary\ body{:}v_i = 0, i = 1, \frac{n}{2}, \tag{9}$$

$$Moving\ body{:}v_j = x_{i-(n/2)} - x_i, i = \frac{n}{2} + 1, \ldots, n. \tag{10}$$

Similarly, like other metaheuristic techniques, velocity cannot be described by the derivative of position regarding time but is stated as displacement in searching space. According to the CB, the velocity afterward the collision can be evaluated

as follows:

$$Stationary\ body{:}v_i' = \frac{(m_{i+(n/2)} + \varepsilon m_{i+(n/2)})v_{i+(n/2)}}{m_i + m_{i+(n/2)}}, \quad i = 1, \ldots \frac{n}{2}, \tag{11}$$

$$Moving\ body{:}v_i' = \frac{(m_i - \varepsilon m_{i-(n/2)})v_i}{m_i + m_{i-(n/2)}}, \quad i = \frac{n}{2}+1, \ldots n, \tag{12}$$

From the expression, ε characterize Coefficient of Restitution, described by the ratio of comparative velocities amongst 2 bodies beforehand and afterward the collisions:

$$\varepsilon = \frac{|v_{i+1}' - v_i'|}{|v_{i+1} - v_i|}. \tag{13}$$

These coefficients vary linearly amongst [0,1] in the optimization algorithm to ensure the balance amongst exploration and exploitation. Afterward the assessment of displacement, it is possible to fix a new position of the stationary and moving bodies:

$$Stationary\ body{:}x_i^{new} = x_j + rand.v_i', i = 1, \frac{n}{2}, \tag{14}$$

$$Moving\ body{:}x_i^{new} = x_{i-(n/2)} + rand.v_i', i = \frac{n}{2}+1, \ldots n, \tag{15}$$

whereas rand denotes a uniform distribution value $[-1, 1]$. This iteration can be executed on all the particles at every iteration and is continued until an ending condition is fulfilled. The entire procedure of the CBO technique has been demonstrated in Algorithm 1.

Algorithm 1 Pseudocode of CBO Algorithm

1. Initializing the CBO population in the searching space
2. Estimating the objective function and determining mass
3. Sort the population for recognizing stationary and moving groups and asses the velocity
4. Estimating the velocity afterwards the collisio
5. The new location can be determined
6. After the Ending criteria are fulfilled, go to step; otherwise, return to step
7. Report the optimum solution
8. Stop

C. PEDESTRIAN DETECTION MODULE

Finally, the DBN model is applied for accurate pedestrian detection. DBN model stacks several restricted Boltzmann machines (RBMs) for the construction of the deep structure. A DBN has a single visible layer and various hidden layers, as demonstrated in Fig.3A. The lower layer forms a directed generative model. But the top two layers form the RBM distribution that can be undirected generative. Thus, L hidden layers $h^{(1)}, h^{(2)}, \ldots, h^{(L)}$ and visible units v are provided, and

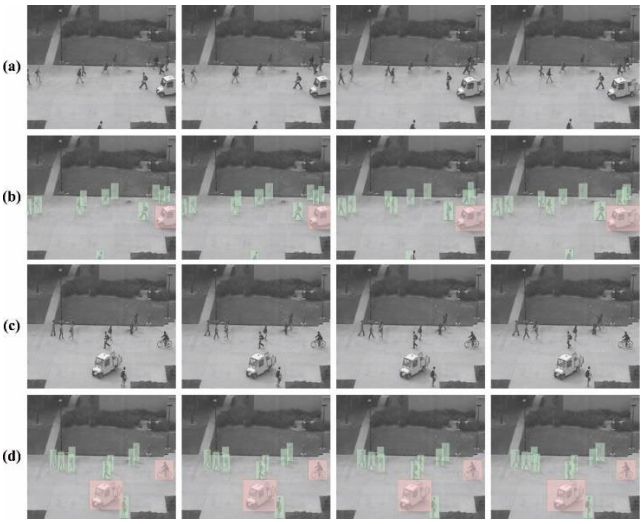

FIGURE 3. Sample visualization results, a) Sequence 1 b) Output of Sequence 1 c) Sequence 2 d) Output of Sequence 2.

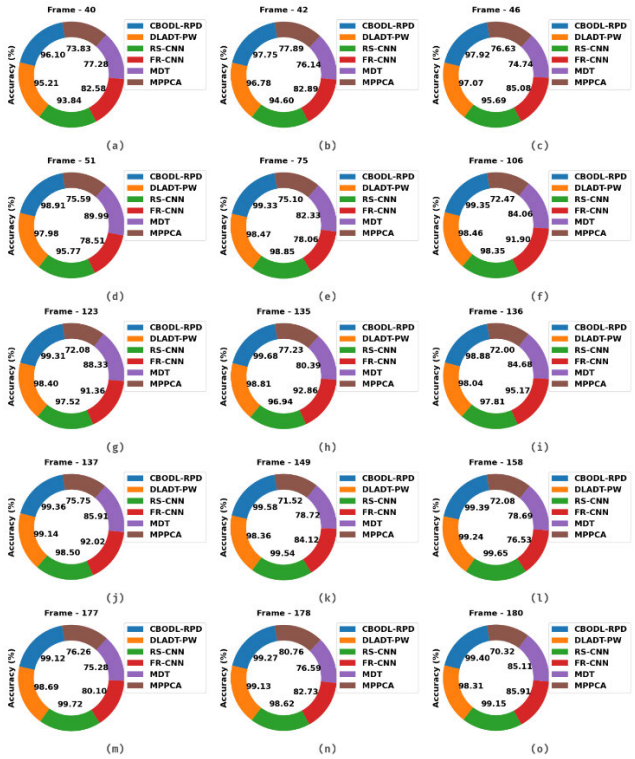

FIGURE 4. Accuy analysis of CBODL-RPD model on Test-004 sequence.

the joint distribution of DBN can be determined below:

$$P\left(v, h^{(1)}, \ldots, h^{(L)}\right) = P\left(v \mid h^{(1)}\right)\left(\prod_{l=1}^{L-2} P\left(h^{(l)} \mid h^{(l+1)}\right)\right) P\left(h^{(L-1)}, h^{(L)}\right) \tag{16}$$

whereas $P(h^{(l)}|h^{(l+1)})$ characterizes conditional distribution for a unit of $l - th$ hidden layers provided unit of $l+1$ hidden

TABLE 1. Overall $accu_y$ of the CBODL-RPD model on test – 004 sequence.

No. of Frames	CBODL-RPD	DLADT-PW	RS-CNN	FR-CNN	MDT	MPPCA
F-40	96.10	95.21	93.84	82.58	77.28	73.83
F-42	97.75	96.78	94.60	82.89	76.14	77.89
F-46	97.92	97.07	95.69	85.08	74.74	76.63
F-51	98.91	97.98	95.77	78.51	89.99	75.59
F-75	99.33	98.47	98.85	78.06	82.33	75.10
F-106	99.35	98.46	98.35	91.90	84.06	72.47
F-123	99.31	98.40	97.52	91.36	88.33	72.08
F-135	99.68	98.81	96.94	92.86	80.39	77.23
F-136	98.88	98.04	97.81	95.17	84.68	72.00
F-137	99.36	99.14	98.50	92.02	85.91	75.75
F-149	99.58	98.36	99.54	84.12	78.72	71.52
F-158	99.39	99.24	99.65	76.53	78.69	72.08
F-177	99.12	98.69	99.72	80.10	75.28	76.26
F-178	99.27	99.13	98.62	82.73	76.59	80.76
F-180	99.40	98.31	99.15	85.91	85.11	70.32

TABLE 2. Overall $accu_y$ of the CBODL-RPD model on test – 007 sequence.

No. of Frames	CBODL-RPD	DLADT-PW	RS-CNN	FR-CNN	MDT	MPPCA
F-78	95.94	94.71	94.58	89.89	84.07	76.14
F-91	99.09	98.22	95.21	91.27	84.90	74.65
F-92	98.93	97.61	98.12	92.74	86.00	65.25
F-110	97.96	96.87	94.52	91.31	83.40	70.86
F-113	94.80	93.53	92.27	87.79	84.37	74.01
F-115	86.79	85.77	83.16	82.61	81.84	70.70
F-125	98.83	97.57	97.25	91.63	88.98	68.47
F-142	98.33	96.92	96.73	94.44	81.69	68.75
F-146	84.43	83.31	83.82	79.93	74.52	78.49
F-147	88.05	86.61	83.00	81.09	80.26	67.17
F-148	75.16	73.64	68.50	54.06	52.86	67.37
F-150	94.25	93.27	87.37	82.87	81.69	70.93
F-178	81.78	80.57	73.53	69.01	62.60	76.14
F-179	79.93	79.00	71.70	63.31	60.52	72.98
F-180	88.13	87.24	83.28	80.41	80.07	77.37

TABLE 3. Overall AVGA of the CBODL-RPD model.

Average Accuracy (%)						
Method	CBODL-RPD	DLADT-PW	RS-CNN	FR-CNN	MDT	MPPCA
Test-004	98.89	98.14	97.64	85.32	81.22	74.63
Test-007	90.83	89.66	86.87	82.16	77.85	71.95

layers, and $P(h^{(L-1)}, h^{(L)})$ correspond the joint distribution of the topmost$L-1$ and L hidden layers.

There are two stages for the training of RBMs: fine-tuning and pre- training. In the pre-training phase, the sDBN was trained via stacking RBM layer-wise to discover the variable space. All the layers are trained as an RBM. Especially, $l-th$ hidden layers are trained as RBMs with the observed data from the output representation of $(l-1)$ hidden layers. This repeat, training all the layers until the top-most layer is reached. Afterward, the pre-training is done, and the finetuning can be implemented for optimizing the network to find optimal parameters. The typical BP and wake-sleep algorithms are proficient at finetuning for discriminative and generative modules, correspondingly. For real-time problems, the attained parameters from the pre-training phase are utilized for initiating DNN; later, deep modules are finetuned via supervised learning algorithms such as *BP*.

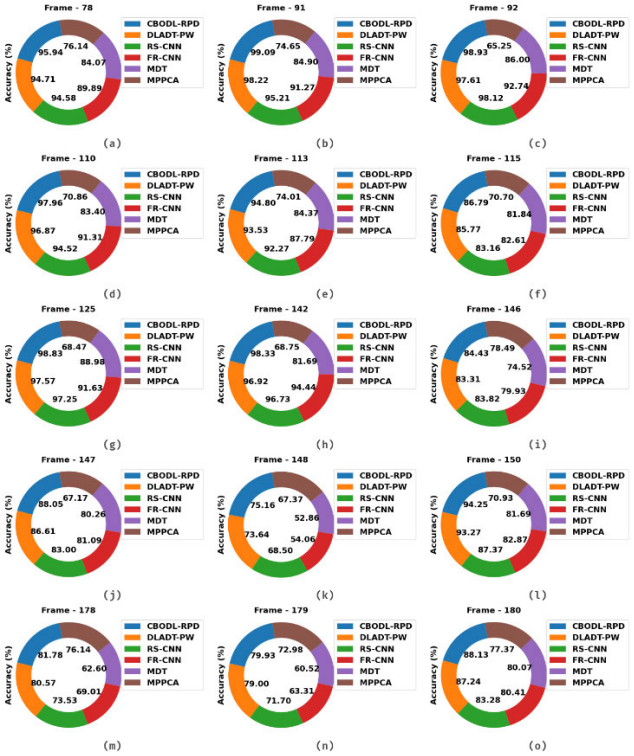

FIGURE 5. $Accu_y$ analysis of CBODL-RPD model on Test-007 sequence.

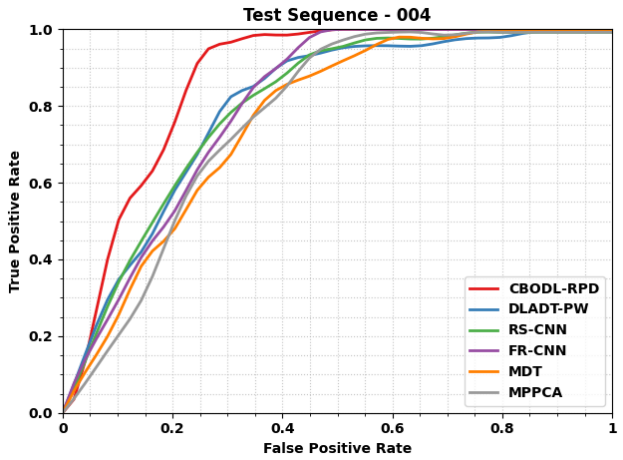

FIGURE 6. TPR of CBODL-RPD model on Test-004 sequence.

IV. RESULTS AND DISCUSSION

The proposed model is simulated using Python 3.6.5 tool on PC i5-8600k, GeForce 1050Ti 4GB, 16GB RAM, 250GB SSD, and 1TB HDD. The parameter settings are given as follows: learning rate: 0.01, dropout: 0.5, batch size: 5, epoch count: 50, and activation: ReLU.

The experimental validation of the CBODL-RPD model is tested using the UCSD Anomaly detection dataset [27]. Fig. 3 demonstrates the sample input sequence and the corresponding tracked sequence by the CBODL-RPD model.

Table 1 and Fig. 4 represent an overall $accu_y$ assessment of the CBODL-RPD method on the test 004 sequence. The figure indicate that the CBODL-RPD algorithm has achieved effectual performance under all frames. For instance, with F-40, the CBODL-RPD model has reached enhanced $accu_y$ of 96.10%. Simultaneously, with F-51, the CBODL-RPD method has attained an enhanced $accu_y$ of 98.91%. Furthermore, with F-106, the CBODL-RPD approach has acquired an enhanced $accu_y$ of 99.35%. Meanwhile, with F-180, the CBODL-RPD technique has reached an enhanced $accu_y$ of 99.40%.

Table 2 and Fig. 5 represent an overall $accu_y$ assessment of the CBODL-RPD model on the test 004 sequence. The figure showed that the CBODL-RPD method had achieved effectual performance under all frames. For instance, with F-78, the CBODL-RPD model has reached an enhanced $accu_y$ of 95.94%. In parallel with F-113, the CBODL-RPD method has

TABLE 4. Comparative TPR study of CBODL-RPD model on test-004 sequence.

True Positive Rate (Test Sequence – 004)						
False Positive Rate	CBODL-RPD	DLADT-PW	RS-CNN	FR-CNN	MDT	MPPCA
0.00	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.05	0.1918	0.1868	0.1666	0.1640	0.1262	0.0959
0.10	0.4953	0.3433	0.3357	0.2903	0.2499	0.1994
0.15	0.6039	0.4342	0.4645	0.4215	0.3988	0.3130
0.20	0.7427	0.5705	0.5831	0.5175	0.4720	0.4872
0.25	0.9239	0.6866	0.6891	0.6462	0.5907	0.6286
0.30	0.9643	0.8154	0.7750	0.7472	0.6614	0.7043
0.35	0.9845	0.8532	0.8305	0.8532	0.7826	0.7750
0.40	0.9845	0.9113	0.8760	0.9113	0.8507	0.8381
0.45	0.9921	0.9315	0.9340	0.9795	0.8785	0.9265
0.50	0.9997	0.9492	0.9517	0.9997	0.9113	0.9668
0.55	0.9997	0.9567	0.9719	0.9997	0.9441	0.9870
0.60	0.9997	0.9567	0.9769	0.9997	0.9769	0.9921
0.65	0.9997	0.9567	0.9744	0.9997	0.9769	0.9921
0.70	0.9997	0.9694	0.9820	0.9997	0.9769	0.9845
0.75	0.9997	0.9769	0.9921	0.9997	0.9921	0.9921
0.80	0.9997	0.9795	0.9921	0.9997	0.9971	0.9921
0.85	0.9997	0.9921	0.9921	0.9997	0.9971	0.9921
0.90	0.9997	0.9946	0.9921	0.9997	0.9971	0.9921
0.95	0.9997	0.9946	0.9921	0.9997	0.9971	0.9921
1.00	0.9997	0.9946	0.9921	0.9997	0.9971	0.9921

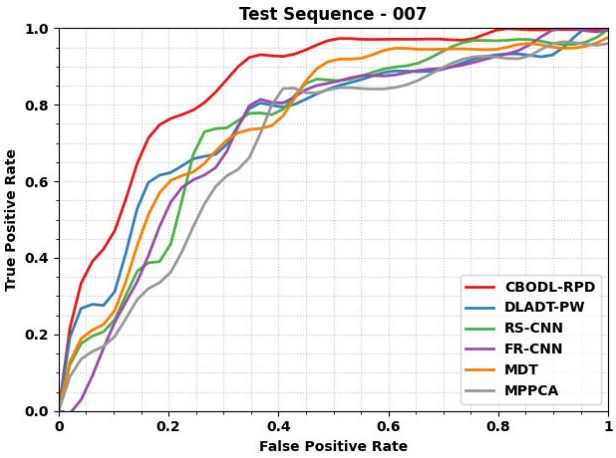

FIGURE 7. TPR of CBODL-RPD model on Test-007 sequence.

achieved an enhanced $accu_y$ of 94.80%. Furthermore, with F-146, the CBODL-RPD approach has reached enhanced $accu_y$ of 84.43%. Simultaneously, with F-180, the CBODL-RPD technique has reached enhanced $accu_y$ of 88.13%.

Table 4 offers an overall average accuracy (AVGA) examination of the CBODL-RPD model. These results indicated that the CBODL-RPD method has improved over other approaches. For instance, on the test-004 sequence, the CBODL-RPD model has gained an $accu_y$ of 98.89%. Contrastingly, the DLADT-PW, RS-CNN, FR-CNN, MDT, and MPPCA algorithms have resulted in decreased $accu_y$ of 98.14%, 97.64%, 85.32%, 81.22%, and 74.63%, respectively. Conversely, on the test-007 sequence, the CBODL-RPD technique has increased $accu_y$ by 90.83%. In Contrast, the DLADT-PW, RS-CNN, FR-CNN, MDT, and MPPCA techniques have resulted in decreased $accu_y$ of 89.66%, 86.87%, 82.16%, 77.85% and 71.95%, correspondingly.

A comparison study of the CBODL-RPD model in terms of TPR on the test-004 sequence is displayed in Table 5 and Fig. 6. The figure demonstrated that the CBODL-RPD method has reached enhanced TPR values under each FPR value. For example, with an FPR of 0.05, the CBODL-RPD approach has a higher TPR of 0.1918 while the DTADT-PW, RS-CNN, FR-CNN, MDT, and MPPCA approaches have demonstrated minimal TPR of 0.1868, 0.1666, 1.1640, 0.1262 and 0.0959 respectively.

Meanwhile, on FPR of 0.50, the CBODL-RPD approach has a higher TPR of 0.9997 while the DTADT-PW, RS-CNN, FR-CNN, MDT, and MPPCA approaches have demonstrated minimal TPR of 0.9492, 0.9517, 0.9997, 0.9113 and 0.9668 correspondingly. Finally, for an FPR of 1.0, the CBODL-RPD approach has a higher TPR of 0.9997 while the DTADT-PW, RS-CNN, FR-CNN, MDT, and MPPCA techniques have demonstrated minimal TPR of 0.9946, 0.9921, 0.9997, 0.9971 and 0.9921 correspondingly.

A brief analysis of the CBODL-RPD approach in terms of TPR on the test-007 sequence is displayed in Table 6 and Fig. 7. The outcomes demonstrated that the CBODL-RPD

TABLE 5. Comparative TPR study of CBODL-RPD model on test-007 sequence.

True Positive Rate (Test Sequence – 007)						
False Positive Rate	CBODL-RPD	DLADT-PW	RS-CNN	FR-CNN	MDT	MPPCA
0.00	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0.05	0.3648	0.2778	0.1877	0.0550	0.2019	0.1464
0.10	0.4643	0.3053	0.2336	0.2255	0.2572	0.1909
0.15	0.6742	0.5602	0.3785	0.3593	0.4630	0.3040
0.20	0.7616	0.6206	0.4211	0.5352	0.5978	0.3552
0.25	0.7900	0.6615	0.6920	0.6072	0.6284	0.4976
0.30	0.8560	0.6852	0.7370	0.6625	0.7005	0.6080
0.35	0.9257	0.7943	0.7789	0.8027	0.7357	0.6693
0.40	0.9260	0.7943	0.7789	0.8027	0.7577	0.8310
0.45	0.9436	0.8146	0.8561	0.8397	0.8626	0.8310
0.50	0.9711	0.8457	0.8631	0.8587	0.9171	0.8425
0.55	0.9711	0.8654	0.8745	0.8757	0.9211	0.8425
0.60	0.9711	0.8866	0.8960	0.8757	0.9453	0.8425
0.65	0.9711	0.8866	0.9067	0.8887	0.9453	0.8611
0.70	0.9711	0.8927	0.9404	0.8957	0.9453	0.8981
0.75	0.9711	0.9167	0.9671	0.9087	0.9453	0.9241
0.80	0.9961	0.9309	0.9671	0.9277	0.9453	0.9241
0.85	0.9961	0.9309	0.9711	0.9502	0.9601	0.9241
0.90	0.9961	0.9309	0.9601	0.9957	0.9502	0.9601
0.95	0.9961	0.9922	0.9601	0.9957	0.9502	0.9601
1.00	0.9961	0.9937	0.9971	0.9957	0.9757	0.9601

TABLE 6. Comparative study of the CBODL-RPD method with recent techniques.

Methods	AUC Score (%)	Computational Time (sec)
CBODL-RPD	95.54	2.10
DLADT-PW	87.24	2.25
RS-CNN	89.23	3.19
FR-CNN	88.08	2.88
MDT	88.08	3.48
MPPCA	87.26	3.69

method has reached enhanced TPR values under each FPR value. For example, with an FPR of 0.05, the CBODL-RPD approach has a maximum TPR of 0.3648 while the DTADT-PW, RS-CNN, FR-CNN, MDT, and MPPCA techniques

have demonstrated minimal TPR of 0.2778, 0.1877, 0.0550, 0.2019 and 0.1464 correspondingly. In the meantime, on FPR of 0.50, the CBODL-RPD method has a higher TPR of 0.9711 while the DTADT-PW, RS-CNN, FR-CNN, MDT, and MPPCA methods have established minimal TPR of 0.8457, 0.8631, 0.8587, 0.9171 and 0.8425 correspondingly. Eventually, for an FPR of 1.0, the CBODL-RPD technique gained a higher TPR of 0.9961 while the DTADT-PW, RS-CNN, FR-CNN, MDT, and MPPCA approaches established minimal TPR of 0.9937, 0.9971, 0.9957, 0.9757 and 0.9601 correspondingly.

Table 6 provides a comprehensive comparative study of the CBODL-RPD approach with recent approaches [28], [29], [30]. The figure revealed that the CBODL-RPD model had shown maximum performance with an increased AUC_{score} of 95.54% and a reduced computation time of 2.10s. These results ensured the betterment of the CBODL-RPD technique.

V. CONCLUSION

In this study, we developed a new CBODL-RPD technique for accurate and efficient pedestrian detection. The core objective of the CBODL-RPD approach is to identify the occurrence of pedestrians and non-pedestrians via the object detection process. For the object detection process, YOLO v4 with Adagrad optimizer is applied. In addition, the CBODL-RPD technique employs the SqueezeNet model to generate feature vectors, and the hyperparameter tuning process is performed via the CBO algorithm. At last, the DBN model is applied for accurate pedestrian detection. A comprehensive experimental analysis is made to demonstrate the significant pedestrian detection results of the CBODL-RPD technique. The comparative outcome reported the improved outcomes of the CBODL-RPD model over other recent models. In future, the presented CBODL-RPD method will be extended for real time pedestrian detection in public places.

ACKNOWLEDGMENT

The authors acknowledge with thanks the Deanship of Scientific Research (DSR), King Abdulaziz University (KAU), Jeddah, Saudi Arabia for technical and financial support.

REFERENCES

[1] A. Brunetti, D. Buongiorno, G. F. Trotta, and V. Bevilacqua, "Computer vision and deep learning techniques for pedestrian detection and tracking: A survey," *Neurocomputing*, vol. 300, pp. 17–33, Jul. 2018.
[2] G. Li, Y. Yang, and X. Qu, "Deep learning approaches on pedestrian detection in hazy weather," *IEEE Trans. Ind. Electron.*, vol. 67, no. 10, pp. 8889–8899, Oct. 2020.
[3] C. Ning, L. Menglu, Y. Hao, S. Xueping, and L. Yunhong, "Survey of pedestrian detection with occlusion," *Complex Intell. Syst.*, vol. 7, no. 1, pp. 577–587, 2021.
[4] L. Chen, S. Lin, X. Lu, D. Cao, H. Wu, C. Guo, C. Liu, and F. Wang, "Deep neural network based vehicle and pedestrian detection for autonomous driving: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, pp. 3234–3246, Jun. 2021.
[5] P. Tumas, A. Nowosielski, and A. Serackis, "Pedestrian detection in severe weather conditions," *IEEE Access*, vol. 8, pp. 62775–62784, 2020.
[6] C. Kyrkou, "YOLOpeds: Efficient real-time single-shot pedestrian detection for smart camera applications," *IET Comput. Vis.*, vol. 14, no. 7, pp. 417–425, 2020.
[7] J. Zhang, L. Lin, J. Zhu, Y. Li, Y. Chen, Y. Hu, and S. C. H. Hoi, "Attribute-aware pedestrian detection in a crowd," *IEEE Trans. Multimedia*, vol. 23, pp. 3085–3097, 2021.
[8] B. Han, Y. Wang, Z. Yang, and X. Gao, "Small-scale pedestrian detection based on deep neural network," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 7, pp. 3046–3055, Jul. 2020.
[9] M. Ragab, "Leveraging mayfly optimization with deep learning for secure remote sensing scene image classification," *Comput. Electr. Eng.*, vol. 108, May 2023, Art. no. 108672.
[10] Z. Yi, S. Yongliang, and Z. Jun, "An improved tiny-YOLOv3 pedestrian detection algorithm," *Optik*, 183, pp. 17–23, Apr. 2019.
[11] X. Song, S. Gao, and C. Chen, "A multispectral feature fusion network for robust pedestrian detection," *Alexandria Eng. J.*, vol. 60, no. 1, pp. 73–85, Feb. 2021.
[12] L. Barba-Guaman, J. E. Naranjo, and A. Ortiz, "Deep learning framework for vehicle and pedestrian detection in rural roads on an embedded GPU," *Electronics*, vol. 9, no. 4, p. 589, 2020.
[13] G. L. Hung, M. S. B. Sahimi, H. Samma, T. A. Almohamad, and B. Lahasan, "Faster R-CNN deep learning model for pedestrian detection from drone images," *Social Netw. Comput. Sci.*, vol. 1, no. 2, pp. 1–9, 2020.
[14] Y. Jiang, G. Tong, H. Yin, and N. Xiong, "A pedestrian detection method based on genetic algorithm for optimize XGBoost training parameters," *IEEE Access*, vol. 7, pp. 118310–118321, 2019.
[15] B. Kim, N. Yuvaraj, K. R. S. Preethaa, R. Santhosh, and A. Sabari, "Enhanced pedestrian detection using optimized deep convolution neural network for smart building surveillance," *Soft Comput.*, vol. 24, no. 22, pp. 17081–17092, 2020.
[16] D. Liu, S. Gao, W. Chi, and D. Fan, "Pedestrian detection algorithm based on improved SSD," *Int. J. Comput. Appl. Technol.*, vol. 65, no. 1, pp. 25–35, 2021.
[17] P. Yang, G. Zhang, L. Wang, L. Xu, Q. Deng, and M. Yang, "A part-aware multi-scale fully convolutional network for pedestrian detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 2, pp. 1125–1137, Feb. 2021.
[18] J. Kolluri and R. Das, "Intelligent multimodal pedestrian detection using hybrid metaheuristic optimization with deep learning model," *Image Vis. Comput.*, vol. 131, Mar. 2023, Art. no. 104628.
[19] B. Abdulrahman and Z. Zhu, "Real-time pedestrian pose estimation, tracking and localization for social distancing," *Mach. Vis. Appl.*, vol. 34, no. 1, p. 8, 2023.
[20] G. Son, J. Kim, and Y. Kim, "Implementation of pedestrian tracking in low-resolution video using multi-camera," *Bull. Netw., Comput., Syst., Softw.*, vol. 12, no. 1, pp. 31–32, 2023.
[21] D. Tang, W. Jin, D. Liu, J. Che, and Y. Yang, "Siam deep feature KCF method and experimental study for pedestrian tracking," *Sensors*, vol. 23, no. 1, p. 482, 2023.
[22] F. Gao, C. Cai, R. Jia, and X. Hu, "Improved YOLOX for pedestrian detection in crowded scenes," *J. Real-Time Image Process.*, vol. 20, no. 2, p. 24, Apr. 2023.
[23] C. J. Lin, S. Y. Jeng, and H. W. Lioa, "A real-time vehicle counting, speed estimation, and classification system based on virtual detection zone and YOLO," *Math. Problems Eng.*, vol. 2021, pp. 1–10, Nov. 2021.
[24] C. Traoré and E. Pauwels, "Sequential convergence of AdaGrad algorithm for smooth convex optimization," *Oper. Res. Lett.*, vol. 49, no. 4, pp. 452–458, 2021.
[25] L. Su, L. Ma, N. Qin, D. Huang, and A. H. Kemp, "Fault diagnosis of high-speed train bogie by residual-squeeze net," *IEEE Trans. Ind. Informat.*, vol. 15, no. 7, pp. 3856–3863, Jul. 2019.
[26] J. Cheng and W. Zhao, "Chaotic enhanced colliding bodies optimization algorithm for structural reliability analysis," *Adv. Struct. Eng.*, vol. 23, no. 3, pp. 438–453, 2020.
[27] *UCSD Anomaly Detection Dataset*. Accessed: Jun. 13, 2023. [Online]. Available: http://www.svcl.ucsd.edu/projects/anomaly/dataset.html
[28] B. Wang and C. Yang, "Video anomaly detection based on convolutional recurrent autoencoder," *Sensors*, vol. 22, no. 12, p. 4647, 2022.
[29] I. V. Pustokhina, D. A. Pustokhin, T. Vaiyapuri, D. Gupta, S. Kumar, and K. Shankar, "An automated deep learning based anomaly detection in pedestrian walkways for vulnerable road users safety," *Saf. Sci.*, vol. 142, Oct. 2021, Art. no. 105356.
[30] M. Ragab and M. Farouk S. Sabir, "Arithmetic optimization with deep learning enabled anomaly detection in smart city," *Comput., Mater. Continua*, vol. 73, no. 1, pp. 381–395, 2022.