Journal of Disability Research

2023 | Volume 2 | Issue 2 | Pages: 49–56
DOI: 10.57197/JDR-2023-0021

Assisting Visually Impaired People Using Deep Learning-based Anomaly Detection in Pedestrian Walkways for Intelligent Transportation Systems on Remote Sensing Images

Hadeel Alsolai[1], Fahd N. Al-Wesabi[2,*], Abdelwahed Motwakel[3] and Suhanda Drar[3]

[1]Department of Information Systems, College of Computer and Information Sciences, Princess Nourah Bint Abdul Rahman University, Riyadh 11671, Saudi Arabia
[2]Department of Computer Science, College of Science and Art at Mahayil, King Khalid University, Abha, Saudi Arabia
[3]Department of Information Systems, College of business administration in Hawtat bani Tamim, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia

Correspondence to:
Fahd N. Al-Wesabi*, e-mail: fwesabi@gmail.com

Received: May 24 2023; Revised: July 19 2023; Accepted: July 20 2023; Published Online: August 01 2023

ABSTRACT

Anomaly detection in pedestrian walkways of visually impaired people (VIP) is a vital research area that utilizes remote sensing and aids to optimize pedestrian traffic and improve flow. Researchers and engineers can formulate effective tools and methods with the power of machine learning (ML) and computer vision (CV) to identifying anomalies (i.e. vehicles) and mitigate potential safety hazards in pedestrian walkways. With recent advancements in ML and deep learning (DL) areas, authors have found that the image recognition problem ought to be devised as a two-class classification problem. Therefore, this manuscript presents a new sine cosine algorithm with deep learning-based anomaly detection in pedestrian walkways (SCADL-ADPW) algorithm. The proposed SCADL-ADPW technique identifies the presence of anomalies in the pedestrian walkways on remote sensing images. The SCADL-ADPW techniques focus on the identification and classification of anomalies, i.e. vehicles in the pedestrian walkways of VIP. To accomplish this, the SCADL-ADPW technique uses the VGG-16 model for feature vector generation. In addition, the SCA approach is designed for the optimal hyperparameter tuning process. For anomaly detection, the long short-term memory (LSTM) method can be exploited. The experimental results of the SCADL-ADPW technique are studied on the UCSD anomaly detection dataset. The comparative outcomes stated the improved anomaly detection results of the SCADL-ADPW technique.

KEYWORDS

anomaly detection, visually impaired people, intelligent transportation systems, deep learning, remote sensing images

INTRODUCTION

Currently, in metropolitan areas, intelligent transportation systems were devised to help decrease traffic, the rate of injuries, accidents, and deaths, diminish environmental pollution, and reduce fuel consumption, etc. (Khan et al., 2022). Such structures use diverse technologies [deep learning (DL), Internet of Things, image processing, and neural network (NN)] for many applications. Alternatively, technology companies (namely Tesla and Google) are attempting to develop self-driving cars that could offer safe travel for persons while the driver is sleepy, which could save his life along with other travelers' lives, with automatic control cars to avoid accidents (Pustokhina et al., 2021). Such automobiles should be fortified with sensors for sensing the atmosphere, finding objects close to the car, and notifying the driver, and have actuators to do real-time processes while the driver is sleepy or fails to pay attention to hazards to avoid accidents (Mansour et al., 2021). Intelligent transportation systems were advanced for real-time operations and automatic road management in the situation of a natural accident (like an icy road, a mountain fall, and an avalanche) or unnatural disasters (namely car traffic accidents and road repairs) (Chang et al., 2022). Such systems utilize sensors as a tool to understand and find the state of travel atmosphere. The data gathered by such sensors by means of communication technologies [dedicated short-range communication (DSRC and WiFi)] to other vehicles on the route are transferred to control centers (Khan et al., 2022).

The captured visual data have details that are precise when compared with substitute data sources namely radar signals, global positioning system (GPS), mobile transmission, etc. (Avola et al., 2022). In addition, it serves a main role in predicting accidents and other abnormal actions by collecting particulars concerning road traffic circumstances. The usage of closed-circuit television (CCTV) is useful in different real-time applications like accidental falls and grade crossing trespassing (Sabuhi et al., 2021). Several computer vision (CV) relies on works were modeled by focusing processes namely behavioral learning, data acquisition, and feature extraction. The intention of studies is to calculate processes like vehicle prediction and observation, scene detection, anomaly prediction approaches, activity examination, multi camera-relied schemes and difficulties, video processing models, traffic observation, and human behavior learning (Dey et al., 2019). Now, anomalous prediction is a sub-field of behavior learning from captured visual scenes. Machine learning (ML) refers to a domain of artificial intelligence (AI) that utilizes statistical approaches to learn hidden patterns from prevailing data and to make decision regarding unobserved records (Abdullah and Jalal, 2023). The core task of machine learners is constructing a general model on the possible dispersal of training samples and generalizing experience to unseen instances. The learning procedure relies upon the data quality exhibited (Aldayri and Albattah, 2022). An instance is presented in a dataset with various properties. Inappropriately, mining proficient features seems to be hard for a few tasks.

This manuscript presents a new sine cosine algorithm with deep learning-based anomaly detection in pedestrian walkways (SCADL-ADPW) method. The SCADL-ADPW techniques focus on the identification and classification of anomalies, i.e. vehicles in the pedestrian walkways of visually impaired people (VIP). To accomplish this, the SCADL-ADPW technique uses the VGG-16 model for feature vector generation. In addition, the SCA approach is designed for the optimal hyperparameter tuning process. For anomaly detection, long short-term memory (LSTM) method can be exploited. The experimental results of the SCADL-ADPW technique are studied on the UCSD anomaly detection dataset.

RELATED WORKS

Yang et al. (2018) devised to seize convolutional neural network (CNN)-related per-pixel semantic segmenters to cover navigational perception necessities in a unified way. It is not critical to observe crosswalk location or traffic light signals, but analyze the states of other vehicles and pedestrians. At the centroid of unification proposal is a DL structure, aspired to achieve robust and potential semantic understanding. Shimakawa et al. (2019) deal with obstacle recognition to support VIP walks. Previous studies had adopted a method of utilizing some image processing methods for finding the steps of staircase on depth imageries by an RGB-D camera. Still, it was hard to carry the RGB-D camera. This study implemented CNN, a type of DL approach to hindrance detection, and advanced an application program functioning on smartphones.

Lin et al. (2018) devised a visual localizer that has global optimization and ConvNet descriptor for attaining strong visual localization for aided navigation. Through concatenation of two compressed convolution layers of GoogLeNet, thousands of bytes are utilized for denoting image proficiently. Hsieh et al. (2021) developed an assistive mechanism to help VIP walk outdoors. Depending on the depth map acquired by ZED 2 and CNN-NN-FAST-SCNN, the image of the environment in front of the VIP can be devised into seven equal divisions.

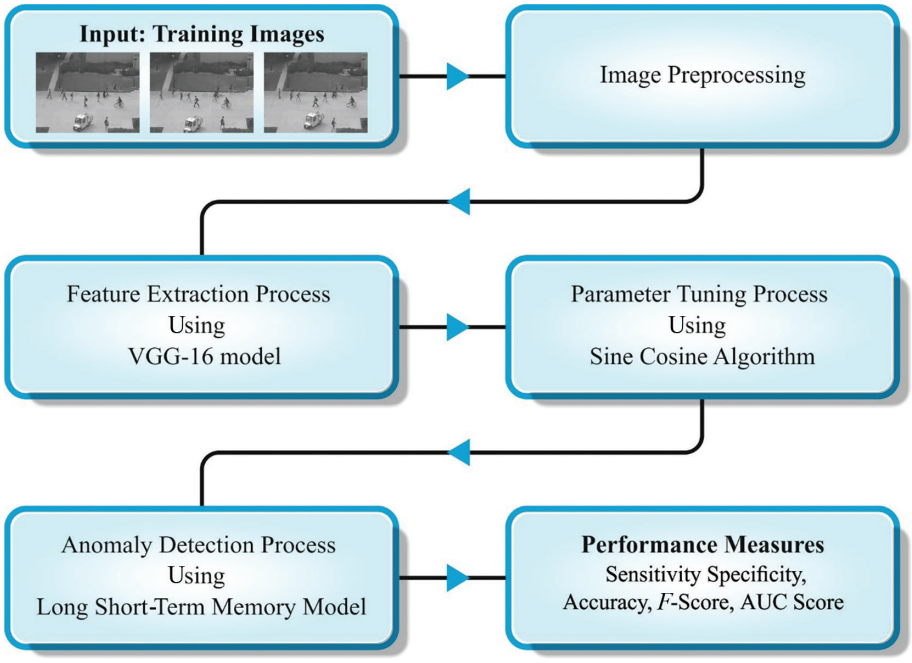

Figure 1: Workflow of the SCADL-ADPW approach. Abbreviation: SCADL-ADPW, sine cosine algorithm with deep learning-based anomaly detection in pedestrian walkways.

Bhattacharya and Asari (2021) devised a method that incorporated wearable devices and Android devices to support VIP and make it simpler to navigate safely. The glove sync and the Android device are connected through Bluetooth connection, and the vibration actuator needs haptic feedback from the user based on the instruction from the paired devices. Le et al. (2022) recommend a novel BGN for camera-related recognition of pedestrian lanes in unstructured scene. For the VIP safety, along with an output segmentation map, network provided two complete-resolution maps of epistemic and aleatoric uncertainty in addition to well-calibrated confidence measure.

PROPOSED ANOMALY DETECTION MODEL

In this manuscript, we have devised a new SCADL-ADPW technique for anomaly detection purposes. The proposed SCADL-ADPW technique identifies the occurrence of anomalies in the pedestrian walkways on remote sensing images. The SCADL-ADPW techniques focus on the identification and classification of anomalies, i.e. vehicles in the pedestrian walkways of VIP. Figure 1 represents the workflow of the SCADL-ADPW algorithm.

Feature extraction module

The SCADL-ADPW techniques use the VGG-16 model for feature vector generation. The VGG-16 is the essential basis of CNN (Deepa and Chokkalingam, 2022). VGG designs and trains object detection where the most important CNN is VGG-16. In 2015, Google researchers introduce the InceptionV3 network. The high computation cost related to CNN can be overcome by the pretrained structure namely VGG-16. Figure 2 illustrates the infrastructure of VGG-16. In the presented method, the pooling, convolution, output and fully connected layers are the four main elements included in CNNs.

Convolutional layer

The building block in CNNs are a convolution layer. The convolution-related method convert inputs into outputs. The fixed height and width with plane composes the input volume. The depth D can be determined by number of planes. The amount of M planes with the volume structures presented in the convolution layer. The amount of filters or kernels with a small size (K) determines all the planes. The weight with filter size is 3 × 3. Y_K shows the matrix value in terms of k^{th} input plane.

$$X_l = F\left(b_l + \sum_{k=1}^{D} w_{kl} * y_k\right); l = 1,\ldots,M, \quad (1)$$

where w_{kl} denotes the parametric weight. Over kth input planes, the 2D matrix $w_{kl} * Y_k$ is generated by means of convolution of these filters. The F activation function transform the outcome and bias function b_l.

The feature map resultant matrix X_l. The feature map denotes M. In this work, the activation function was ReLU. While keeping all the positive input, it changes every value of input to 0.

$$F(y) = \text{Max}(0, y) \quad (2)$$

The convolution layer output employs the ReLU function. The network learning speed and classification accuracy improved.

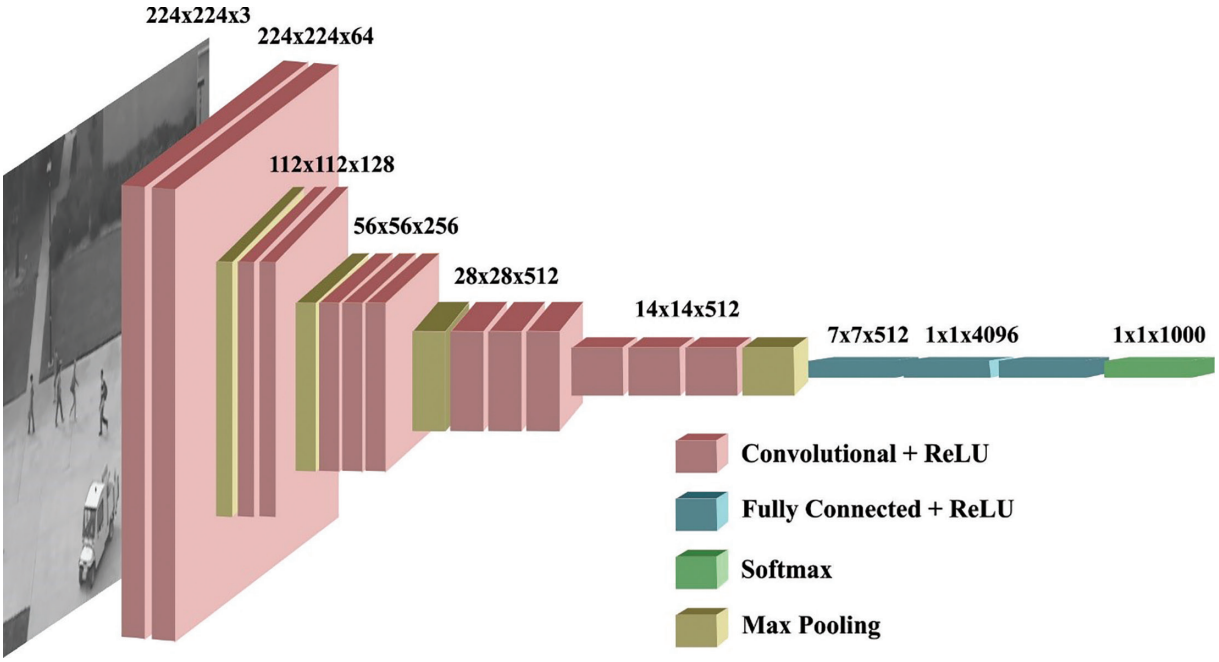

Figure 2: Structure of VGG16.

Pooling layer

The feature map size can be decreased by means of the pooling layer after all the blocks of a layer or convolution layer. The crucial data are preserved while avoiding inappropriate details. The squared section and the stride (s) of feature map size (F) are the two crucial parameter to determine pooling.

The $s = 2$ and $F = 2$ with five pooling layers existing in VGG-16. The block organizes 13 convolution layers which comprise the VGG-16.

Fully connected and output layer

Attain feature map set afterwards the successive and pooling convolution layer. The input–output function implemented by all the neurons of the fully connected (FC) layer is represented as follows:

$$Z = F\left(b + \sum_{j}^{n} w_j Y_j\right) \tag{3}$$

Therefore, the w_j and Y_j neuron denote input features and neuron weight. The FC layer output is Z:

$$F(x_j) = Softmax(x_l) = \frac{e^{(x_l)}}{\sum_{k=1}^{8} e^{(x_k)}}, for\, l = 1,\ldots..,8 \tag{4}$$

The probability distribution outcome can be attained by using the output layer of all the node's outcomes.

Hyperparameter tuning module

Here, the SCA approach is designed for the optimal hyperparameter tuning process. SCA is a population-based optimization technique which creates several initial choices and enables them to fluctuate toward optimum solution through the sin and cos operators (Pham and Nguyen, 2023). The optimization method of population-related optimization algorithm split into two stages: exploration and exploitation.

The exploitation and exploration stages are as follows:

$$X_i^{t+1} = X_i^t + r_1 * \sin(r_2) * |r_3 * P - X_i^t|, \tag{5}$$

$$X_i^{t+1} = X_i^t + r_1 * \cos(r_2) * |r_3 * P - X_i^t|. \tag{6}$$

$$X_i^{t+1} = \begin{cases} X_i^t + r_1 * \sin(r_2) * |r_3 * P - X_i^t|; r_4 \le 0..5 \\ X_i^t + r_1 * \cos(r_2) * |r_3 * P - X_i^t|; r_4 \ge 05 \end{cases}, \tag{7}$$

where r_4 shows the random value in [0,1].

The key parameters are r_1, r_2, r_3, and r_4.

- r_1 determines various directions within space between the solution and the space outside the solution or the destination.
- r_2 provides distance to travel toward or outward the destination.
- r_3 shows random weight to emphasize weight of the exploitation ($r_3 < 1$) or exploration ($r_3 > 1$).

Effect of *sin* and *cos* functions in Eqs. (5) and (6). Eqs. (5) and (6) determine the spatial area between two solutions in the search range. But, the search space was expanded to high dimension. The cyclical model of sin and cos functions enables relocating around various solutions. This method guarantees exploitation of space determined among two solutions.

Furthermore, to avoid local optima, solution should be searched outside the space between the solution and the destination.

The range of *sin* and *cos* functions in Eqs. (5) to (6) is changed to balance the exploration and exploitation stages:

$$r_1 = a - t\frac{a}{T}, \tag{8}$$

In Eq. (8), t denotes the existing iteration, and T shows the maximum amount of iterations.

Anomaly detection module

For anomaly detection, the LSTM model is utilized. The key features of LSTM-NN involve the usage of gating modules to optimize the problem of RNN being incapable of transmitting prior data for a longer time, resolve the problems of gradient exploding and vanishing at longer sequence training, and accomplish long-term time series learning (Hong and Tian, 2023). It includes forgetting, input, and output gates that could selectively learn input data and store relevant data in the storage unit. LSTM is often employed in different domains like speech recognition, unmanned driving, intelligent robots, etc.

$$tanh\ (x) = \frac{\sinh(x)}{\cosh(x)} \tag{9}$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{10}$$

$$h_t = o_t \times tanh\ (C) \tag{11}$$

$$C_t = f_t \times C_{t-1} + i_t \times c_t \tag{12}$$

$$f_t = \sigma(w_f \cdot [h_{t-1}, X_t] + b_f) \tag{13}$$

$$i_t = \sigma(w_i \cdot [h_{t-1}, X_t] + b_i) \tag{14}$$

$$o_t = \sigma(w_o \cdot [h_{t-1}, X_t] + b_0), \tag{15}$$

where C_{t-1} and C_t correspondingly signify the cell state at t–1 and t time, c_t denotes the cell status update value at t time, h_{t-1} and h_t correspondingly characterize the output at t–1 and t time, 0_t, i_t and f_t signify output, input, and forgetting gates, correspondingly; X_t shows the input at t time, 0_t, i_t and f_t denotes each matrix within [0,1] *tanh* and σ denotes activation functions. σ shows the sigmoid function that is S-type function could map the output of real number. The *tanh* function ranges within (–1,1). w_0, w_i, and w_f correspondingly signify the weight matrices of the three gates; b_0, b_i, and b_f signify the bias terms of the three gating functions.

Based on the abovementioned derived equations and LSTM structure diagram, the input of LSTM-NN at t time can be defined using the input X_t at the existing time and output h_{t-1} at $t-1$ time. The cell status update value c_t at t time can be defined using i_t input gate that can be used for selecting relevant data for updating unit state C_t. The forget gate f_t is used for controlling cell state at $t-1$ time and filter storage data.

The joint effect of f_t and i_t forgetting and input gates is to forget worthless data and update relevant data for obtaining unit state C_t at time t. The output gate 0_t can be leveraged for controlling the cell state C_t at t time. Initially, the *tanh* function maps the value of the C_t cell state at (–1,1); then the input gate o_t analyzes the output h_t at t time.

EXPERIMENTAL EVALUATION

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.

In this section, the anomaly recognition in pedestrian walkways results is examined on the UCSD anomaly detection dataset. Figure 3 illustrates the sample images.

In Table 1 and Figure 4, the recognition accuracy of the SCADL-ADPW approach is examined with current methods on the Test004 dataset. The outcomes guaranteed the improvement of the SCADL-ADPW technique on the detection process. S-40 sequence had an increased detection accuracy of 99.11%, while the DLADT-PW, RS-CNN, Fast R-CNN, and MDT methods have reported decreased detection accuracy of 96, 92.51, 80.41, and 78.21% respectively. Likewise, on S-46 sequence, an increased detection accuracy of 99.85% was observed, while the DLADT-PW, RS-CNN, Fast R-CNN, and MDT methodologies have reported decreased detection accuracy of 98.14, 96.60, 85.17, and 75.48% respectively. Similarly, on S-106

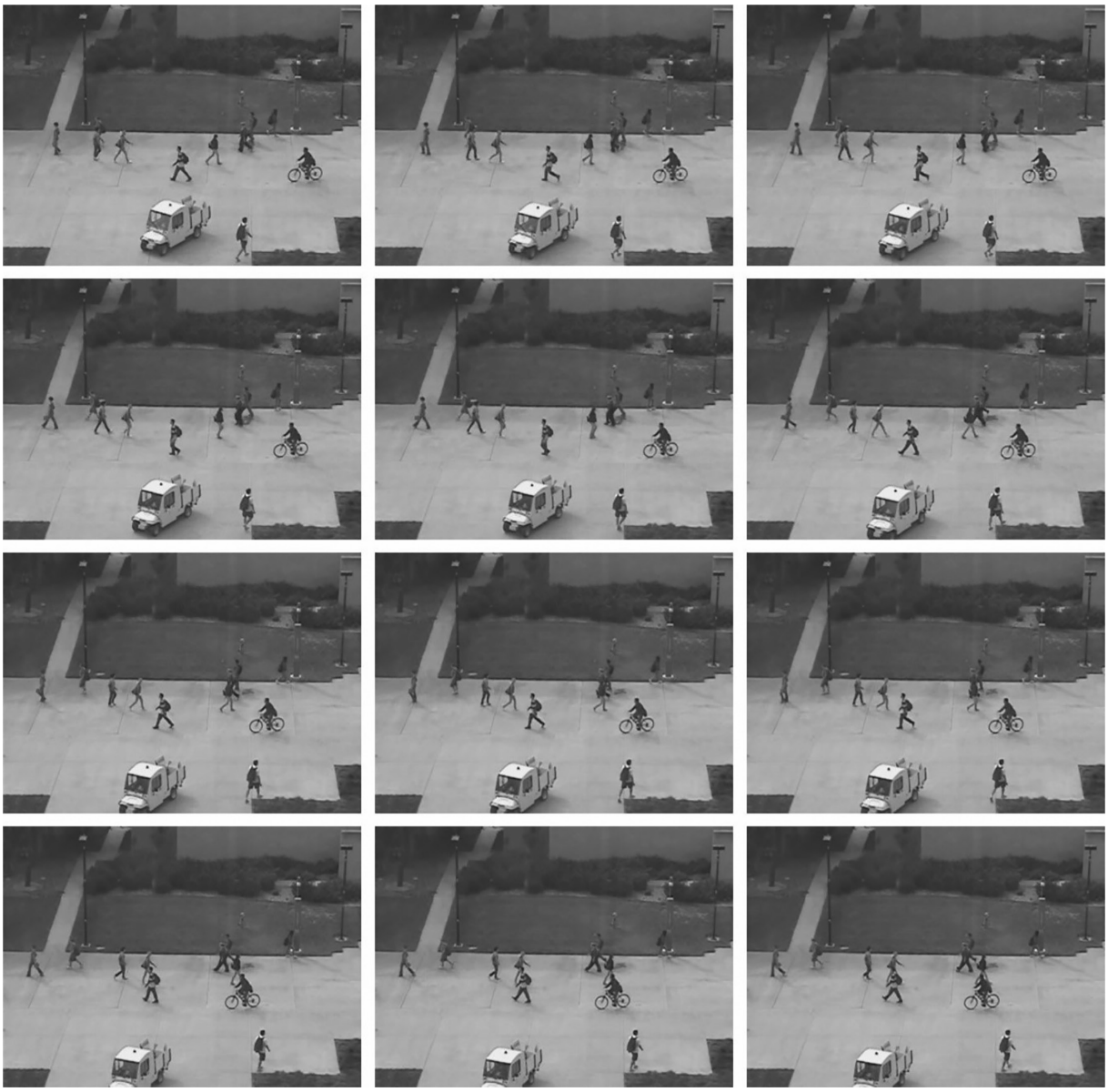

Figure 3: Sample images.

Table 1: Recognition accuracy of the SCADL-ADPW approach with other systems on Test004 dataset.

Test004 Sequences	DLADT-PW	RS-CNN	Fast R-CNN	MDT	SCADL-ADPW
S-40	96.00	92.51	80.41	78.21	99.11
S-42	97.31	94.10	81.55	77.17	99.57
S-46	98.14	96.60	85.17	75.48	99.85
S-51	98.18	96.39	80.88	90.68	99.44
S-75	99.81	99.74	76.92	81.66	99.15
S-106	97.62	100.55	90.77	82.73	99.23
S-123	96.52	97.93	90.97	87.77	99.32
S-135	99.67	99.16	91.05	81.93	99.78
S-136	96.92	97.61	93.98	83.39	99.52
S-137	98.04	99.49	90.81	84.76	99.37
S-149	98.79	99.42	84.06	80.15	99.61
S-158	98.78	99.25	78.58	78.54	99.77
S-177	97.87	98.12	79.05	76.71	98.91
S-178	98.07	100.12	82.81	76.59	99.16
S-180	99.39	99.42	85.96	84.24	99.10
Average	98.07	98.03	84.86	81.33	99.24

Abbreviation: SCADL-ADPW, sine cosine algorithm with deep learning-based anomaly detection in pedestrian walkways.

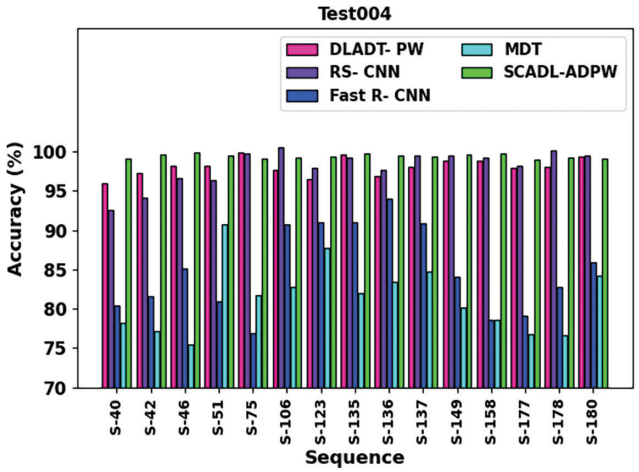

Figure 4: *Accu*$_y$ outcome of the SCADL-ADPW approach on Test004 dataset. Abbreviation: SCADL-ADPW, sine cosine algorithm with deep learning-based anomaly detection in pedestrian walkways.

sequence, an increased detection accuracy of 99.23% was observed, while the DLADT-PW, RS-CNN, Fast R-CNN, and MDT methodologies have reported decreased detection accuracy of 97.62, 100.55, 90.77, and 82.73% respectively. Eventually, S-137 sequence had an increased detection accuracy of 99.37%, while the DLADT-PW, RS-CNN, Fast R-CNN, and MDT methods have reported decreased detection accuracy of 98.04, 99.49, 90.81, and 84.76%, respectively. Finally, on S-180 sequence, an increased detection accuracy of 99.10% was observed, while the DLADT-PW, RS-CNN, Fast R-CNN, and MDT methods had decreased detection accuracy of 99.39, 99.42, 85.96, and 84.24%, respectively.

In Table 2 and Figure 5, the recognition accuracy of the SCADL-ADPW approach is inspected with current approaches on the Test004 dataset. The results guaranteed the improvement of the SCADL-ADPW technique on the detection process. On S-40 sequence, an increased detection accuracy of 96.78% was observed, while the DLADT-PW,

Table 2: Recognition accuracy of the SCADL-ADPW approach with other systems on Test007 dataset.

Test007 Sequence	DLADT-PW	RS-CNN	Fast R-CNN	MDT	SCADL-ADPW
S-40	99.36	97.44	92.26	87.73	96.78
S-42	99.15	98.77	95.21	88.18	96.14
S-46	99.89	99.91	95.76	89.31	96.97
S-51	99.02	98.32	94.59	85.60	96.78
S-75	96.59	96.01	91.31	87.57	96.32
S-106	89.42	86.97	85.98	85.46	96.35
S-123	99.16	99.44	95.08	91.38	96.65
S-135	99.13	98.67	97.13	84.64	96.18
S-136	86.85	86.68	82.49	77.64	96.77
S-137	88.96	85.52	84.48	83.00	96.80
S-149	76.64	70.81	58.48	56.15	96.93
S-158	96.81	90.84	85.80	84.07	96.15
S-177	82.82	76.23	72.65	64.98	96.95
S-178	81.53	75.88	66.15	64.45	96.65
S-180	90.05	86.15	83.47	83.78	96.47
Average	92.36	89.84	85.39	80.93	96.59

Abbreviation: SCADL-ADPW, sine cosine algorithm with deep learning-based anomaly detection in pedestrian walkways.

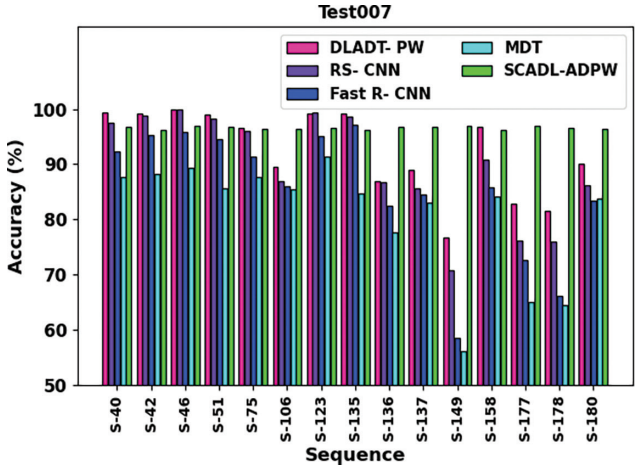

Figure 5: *Accu*$_y$ outcome of the SCADL-ADPW approach on Test007 dataset. Abbreviation: SCADL-ADPW, sine cosine algorithm with deep learning-based anomaly detection in pedestrian walkways.

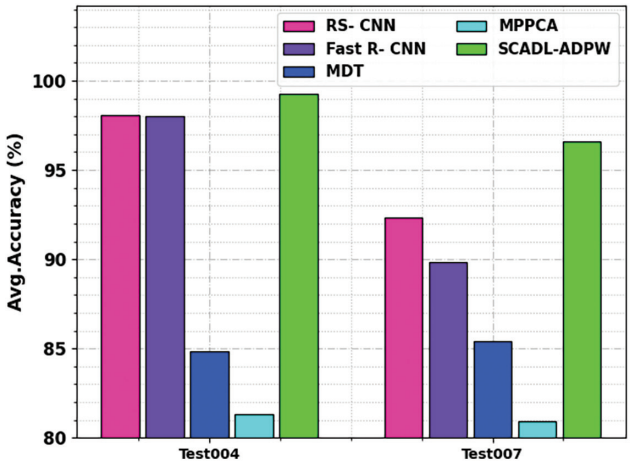

Figure 6: Average *accu*$_y$ outcome of the SCADL-ADPW approach on two datasets. Abbreviation: SCADL-ADPW, sine cosine algorithm with deep learning-based anomaly detection in pedestrian walkways.

RS-CNN, Fast R-CNN, and MDT methods had decreased detection accuracy of 99.36, 97.44, 92.26, and 87.73%, respectively. Similarly, on S-46 sequence, an increased detection accuracy of 96.97% was observed, while the DLADT-PW, RS-CNN, Fast R-CNN, and MDT methodologies have reported decreased detection accuracy of 99.89, 99.91, 95.76, and 89.31%, correspondingly. Likewise, on S-106 sequence, an increased detection accuracy of 96.35% was observed, while the DLADT-PW, RS-CNN, Fast R-CNN, and MDT methods had decreased detection accuracy of 89.42, 86.97, 85.98, and 85.46%, correspondingly. Eventually, S-137 sequence had an increased detection accuracy of 96.80%, while the DLADT-PW, RS-CNN, Fast R-CNN, and MDT algorithms had decreased detection accuracy of 88.96, 85.52, 84.48, and 83%, respectively. Finally, S-180 sequence had an increased detection accuracy of 96.47%, while the DLADT-PW, RS-CNN, Fast R-CNN, and MDT systems had decreased detection accuracy of 90.05, 86.15, 83.47, and 83.78%, correspondingly.

The overall anomaly detection results of the SCADL-ADPW technique have been compared with other DL models in Table 3 and Figure 6 (Pustokhina et al., 2021). The outcomes notify the improved average accuracy results over other models. On the Test004 dataset, the SCADL-ADPW technique achieves improved average *accu*$_y$ of 99.24% while the RS-CNN, Fast R-CNN, MDT, and MPPCA techniques attain decreased average *accu*$_y$ of 98.07, 98.03, 84.86, and 81.33%, respectively. Besides, on the Test007 dataset, the SCADL-ADPW method achieves improved average *accu*$_y$ of 96.59%, while the RS-CNN, Fast R-CNN, MDT, and MPPCA approaches decreased average *accu*$_y$ of 92.36, 89.84, 85.39, and 80.93%, respectively.

These results ensured the effectual anomaly detection performance of the SCADL-ADPW technique.

Table 3: Average accuracy of the SCADL-ADPW approach with other systems on two datasets.

Techniques	RS-CNN	Fast R-CNN	MDT	MPPCA	SCADL-ADPW
Test004	98.07	98.03	84.86	81.33	99.24
Test007	92.36	89.84	85.39	80.93	96.59

Abbreviation: SCADL-ADPW, sine cosine algorithm with deep learning-based anomaly detection in pedestrian walkways.

CONCLUSION

In this manuscript, we have designed a new SCADL-ADPW technique for anomaly detection purposes. The proposed SCADL-ADPW technique identifies the occurrence of anomalies in the pedestrian walkways on remote sensing images. The SCADL-ADPW techniques focus on the identification and classification of anomalies, i.e. vehicles in the pedestrian walkways of VIP. To accomplish this, the SCADL-ADPW technique uses the VGG-16 model for feature vector generation. In addition, the SCA approach is designed for the optimal hyperparameter tuning process. For anomaly detection, the LSTM model can be exploited. The experimental results of the SCADL-ADPW technique are studied on the UCSD anomaly detection dataset. The comparative outcomes stated the improved anomaly detection results of the SCADL-ADPW technique.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest in association with the present study.

ACKNOWLEDGMENT

The authors extend their appreciation to the King Salman Center for Disability Research for funding this work through Research Group no KSRG-2023-334.

REFERENCES

Abdullah F. and Jalal A. (2023). Semantic segmentation based crowd tracking and anomaly detection via neuro-fuzzy classifier in smart surveillance system. *Arab. J. Sci. Eng.*, 48(2), 2173-2190.

Aldayri A. and Albattah W. (2022). Taxonomy of anomaly detection techniques in crowd scenes. *Sensors,* 22(16), 6080.

Avola D., Cannistraci I., Cascio M., Cinque L., Diko A., Fagioli, A., et al. (2022). A novel GAN-based anomaly detection and localization method for aerial video surveillance at low altitude. *Remote Sens.*, 14(16), 4110.

Bhattacharya A. and Asari V.K. (2021). Wearable walking aid system to assist visually impaired persons to navigate sidewalks. In: *2021 IEEE Applied Imagery Pattern Recognition Workshop (AIPR),* IEEE, Washington, DC, USA. pp. 1-7.

Chang Y., Tu Z., Xie W., Luo B., Zhang S., Sui H., et al. (2022). Video anomaly detection with spatio-temporal dissociation. *Pattern Recognit.,* 122, 108213.

Deepa N. and Chokkalingam S.P. (2022). Optimization of VGG16 utilizing the arithmetic optimization algorithm for early detection of Alzheimer's disease. *Biomed. Signal Process. Control,* 74, 103455.

Dey A., Mohammad F., Ahmed S., Sharif R. and Saif A.S. (2019). Anomaly detection in crowded scene by pedestrians behaviour extraction using long short term method: a comprehensive study. *Int. J. Educ. Manag. Eng.*, 9(1), 51.

Hong J. and Tian W. (2023). Prediction in catalytic cracking process based on swarm intelligence algorithm optimization of LSTM. *Processes*, 11(5), 1454.

Hsieh I., Cheng H.C., Ke H.H., Chen H.C. and Wang W.J. (2021). A CNN-based wearable assistive system for visually impaired people walking outdoors. *Appl. Sci.,* 11(21), 10026.

Khan A.A., Nauman M.A., Shoaib M., Jahangir R., Alroobaea R., Alsafyani M., et al. (2022). Crowd anomaly detection in video frames using fine-tuned AlexNet model. *Electronics*, 11(19), 3105.

Le H.T., Phung S.L. and Bouzerdoum A. (2022). Bayesian Gabor network with uncertainty estimation for pedestrian lane detection in assistive navigation. *IEEE Trans. Circuits Syst. Video Technol.,* 32(8), 5331-5345.

Lin S., Cheng R., Wang K. and Yang K. (2018). Visual localizer: outdoor localization based on convnet descriptor and global optimization for visually impaired pedestrians. *Sensors*, 18(8), 2476.

Mansour R.F., Escorcia-Gutierrez J., Gamarra M., Villanueva J.A. and Leal N. (2021). Intelligent video anomaly detection and classification using faster RCNN with deep reinforcement learning model. *Image Vis. Comput.,* 112, 104229.

Pham V.H.S. and Nguyen V.N. (2023). Cement transport vehicle routing with a hybrid sine cosine optimization algorithm. *Adv. Civ. Eng.*, 2023.

Pustokhina I.V., Pustokhin D.A., Vaiyapuri T., Gupta D., Kumar S. and Shankar K. (2021). An automated deep learning based anomaly detection in pedestrian walkways for vulnerable road users safety. *Saf. Sci.,* 142, 105356.

Sabuhi M., Zhou M., Bezemer C.P. and Musilek P. (2021). Applications of generative adversarial networks in anomaly detection: a systematic literature review. *IEEE Access*, 9, 161003-161029.

Shimakawa M., Taguchi I., Okuma C. and Kiyota K. (2019). Smartphone application program of obstacle detection for visually impaired people. *ICIC Express Lett. Part B Appl.*, 10(3), 219-226.

Yang K., Cheng R., Bergasa L.M., Romera E., Wang K. and Long N. (2018). Intersection perception through real-time semantic segmentation to assist navigation of visually impaired pedestrians. In: *2018 IEEE International Conference on Robotics and Biomimetics (ROBIO), December 12–15, 2018,* IEEE, Kuala Lumpur, Malaysia. pp. 1034-1039.