



OPEN Enhanced pedestrian walkway object detection using deep learning and pelican optimization algorithm for assisting disabled persons

Fadwa Alrowais¹, Mona Almofarreh² & Radwa Marzouk^{3,4}✉

Walking is a significant transportation method, but the convenience of pedestrian surroundings for individuals with blindness is highly challenging. Pedestrians with blindness familiarize themselves with guidelines in their surroundings, which might be artificial or natural. To overcome these troubles, it is highly significant for them to perceive the features of an environment. Currently, numerous methods like long white canes and GPS are deployed to improve pedestrian walkways for sightless people. So, they can utilize it as the primary assistive device for recognition and also the vital ecological features for persons with disability. Recently, a growing amount of success has been conveyed for vision navigation tasks depend upon deep learning (DL) and machine learning (ML) networks to aid visually impaired people. This study proposes an Enhanced Pedestrian Walkway Object Detection and Pelican Optimization Algorithm for Assisting Disabled Persons (EPWOD- POAADP) method. The main intention of the EPWOD-POAADP method is to enhance pedestrian walkways for blind people's navigation. At first, the image pre-processing stage applies median filtering (MF) to eliminate the noise in the input data. Furthermore, the Faster R-CNN model is employed for the object detection process to identify and locate objects within an image. The CapsNet model is used for the feature extraction process. In addition, the wavelet neural network (WNN) technique is implemented for the detection and classification process. Finally, the hyperparameter selection of the WNN model is performed using the pelican optimization algorithm (POA) technique. The experimental evaluation of the EPWOD-POAADP approach is examined under the UCSD anomaly detection dataset. The outcomes indicated the enhanced performance of the EPWOD-POAADP approach compared to recent approaches.

Keywords Pedestrian walkway, Object detection, Pelican optimization algorithm, Disabled persons, Faster R-CNN

World health organization (WHO) states that at least one billion individuals will be blind in 2020. It is mainly affected by age-related cataracts, neurological defects from birth, and uncorrected refractive errors¹. For those who are blind, either confidence or independence to undertake everyday living routines was affected². People determined by visual ailments and deficiencies require support to triumph through daily assignments, like exploring and moving to unknown settings. Despite several developments in innovation, blindness endures a significant challenge³. Usually, pedestrians with visual impairments disregard much data about their instant setting that sighted people might take without proof⁴. Whereas multiple experts are recompensing for missed data over improved awareness of other gestures and the utilization of navigational assistance, either lower-tech, for instance, guide dogs or white canes, or higher-tech, for example, GPS gadgets, there are still multiple circumstances in which people with visual impairments are unable to travel individually they would like⁵.

¹Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, 11671 Riyadh, Saudi Arabia. ²Department of Computer Science, College of Computer and Information Sciences, King Saud University, Saudi Arabia, P.O.Box 51178, 11543 Riyadh, Saudi Arabia. ³Department of Mathematics, Faculty of Science, Sudan University, Khartoum, Sudan. ⁴King Salman Center for Disability Research, 11614 Riyadh, Saudi Arabia. ✉email: radwa@sci.cu.edu.eg

For people with visual impairments, moving to a novel setting might be a specific difference in proficiency⁶. Consequently, while travellers with visual impairments search unknown targets, they are frequently required to plan forward widely to memorize and attain directions, and several may seek support from others comprising family members, friends, and specialized trainers to inform themselves of an unknown location⁷. While moving rather known routes, managing sudden necessities in a journey, like finding a drink, food, or toilet, could be challenging. Primarily, every requirement could involve mastery of a further path, and it might not be very easy to predict each route one may need to know to improve⁸. Investigators are aiming at this concern to advance assistants or supportive gadgets for visually impaired people (VIPs). Nowadays, multiple computer vision (CV) depends on jobs modelled by aiming at processes like data acquisition, feature extraction, and behavioural learning⁹. Deep Learning (DL) and Machine Learning (ML) relate to a field of Artificial Intelligence (AI) that employs statistical models to learn unseen patterns from dominant information and to make decisions in terms of unnoticed registers, where DL and ML-based models are effective assistive methods to assist visually impaired walking outdoors and indoors¹⁰.

This study proposes an Enhanced Pedestrian Walkway Object Detection and Pelican Optimization Algorithm for Assisting Disabled Persons (EPWOD-POAADP) method. The main intention of the EPWOD-POAADP method is to enhance pedestrian walkways for blind people's navigation. At first, the image pre-processing stage applies median filtering (MF) to eliminate the noise in the input data. Furthermore, the Faster R-CNN model is employed for the object detection process to identify and locate objects within an image. The CapsNet model is used for the feature extraction process. In addition, the wavelet neural network (WNN) technique is implemented for the detection and classification process. Finally, the hyperparameter selection of the WNN model is performed using the pelican optimization algorithm (POA) technique. The experimental evaluation of the EPWOD-POAADP approach is examined using a benchmark image dataset. The major contribution of the EPWOD-POAADP approach is listed below.

- The EPWOD-POAADP model initially utilizes MF to eliminate impulse noise and preserve edge details, improving image clarity. This enhances the quality of inputs for subsequent processing stages and strengthens the model's overall robustness and reliability.
- The Faster R-CNN method is integrated into the framework to enable precise and efficient object detection by generating accurate region proposals. This ensures high output in identifying relevant targets within the input images. Its inclusion significantly improves the detection performance of the EPWOD-POAADP technique.
- The CapsNet technique is employed for robust feature extraction, effectively preserving the spatial hierarchies and relationships in visual data. This improves the model's capability of comprehending part-whole relationships and orientation discrepancies. Its integration strengthens the feature representation, resulting in enhanced classification result.
- The EPWOD-POAADP approach employs the WNN technique to detect and classify the extracted features, enabling multi-resolution analysis of complex patterns. This improves the method's capacity to capture time and frequency information, improving detection result and reliability.
- The EPWOD-POAADP methodology implements the POA model to tune the WNN's hyperparameters optimally, improving its learning efficiency and convergence speed. This results in enhanced detection and classification result. The utilization of POA ensures robust and reliable model performance.
- Integrating POA-tuned WNN with CapsNet and Faster R-CNN forms a unique hybrid architecture that effectively integrates robust feature extraction, precise object detection, and optimized classification. This synergy enhances the efficiency in detection tasks. The novel method utilizes the strengths of each component, resulting in a robust and optimized solution.

Literature of works

Bhatlawande et al.¹¹ proposed a model for aiding visually impaired people (VIP) by classifying and detecting succeeding difficulties in pedestrians and vehicles on the way. While walking on pathways or roads, VIPs have inadequate admittance to data about their settings; thus, identifying succeeding cars or pedestrians is crucial for their safety. Walking from one position to another is one of the most complicated jobs for VIPs. Trained dogs and white canes are the most frequently employed instruments to help VIPs navigate and travel. Kumar et al.¹² projected an obstacle recognition structure compounding a road object detection method and a road anomaly recognition method, utilizing a parallel process for rapid real-world implementation. These techniques depend upon CNN backbones, utilize TL, and are skilled in custom datasets gathered physically in unorganized settings. Adi et al.¹³ intended to design and evaluate disability-friendly pedestrian pathways for safety and optimum availability in Indonesia. A pedestrian pathways technique was obtained by utilizing a data model triangulation. In¹⁴, a new approach to determining the ground impedance under only a single shoe is projected in this manuscript. These models utilize bipolar electrodes to terminate the leakage existing from the body. A finite element analysis (FEA) methodology is implemented to exhibit the bi-polar electrode benefits through unipolar electrodes. A laboratory, testing area, and error analysis are accomplished on the advanced prototype to see the method's utility.

Hamadi and Latoui¹⁵ introduced an innovative and favourable indoor localization solution to address the restrictions of either SLAM or PDR over their synergistic incorporation. In reality, to precise the cumulative errors of the developed localization method and consequently enhance the precision. Yoshikawa and Premachandra¹⁶ projected an automated sensing model for pedestrian crossings that employs images from cameras connected to them. The developed model keeps unique features, allowing it to manage difficult circumstances that conventional models contend with effectively. It outperforms identifying crosswalks even in low-light circumstances at night, while illumination stages might differ. Guo and Shen¹⁷ intended to utilize the Internet of Things (IoT) and other smart gadgets to advance a smart pedestrian crossing that is safer and more beneficial, specifically for the

visually impaired and for movement. IoT and other smart gadgets were primarily designated to alert drivers and assist pedestrians effectively. Then, the indication model of the LED light and the rapid process of hearable pedestrian bollards were reshaped to enhance their effectiveness for movement and support of visually impaired individuals. Moreover, virtual reality (VR) was utilized to assume the smart pedestrian crossing. Eventually, a smart space design concept for the smart pedestrian crossing is presented.

The limitations of the existing studies comprise the lack of large-scale real-time testing in dynamic environments and minimal consideration for adaptability across varied terrains or unstructured settings. Most approaches depend on static sensors or specific hardware configurations, mitigating flexibility and scalability. The dependency on CNN and TL models often needs extensive computational resources, which may not suit portable devices. The utilization of FEA and VR is limited to simulation environments with minimal real-world validation. A major research gap is in integrating lightweight, real-time, cross-environment pedestrian support systems for VIPs that merge IoT, adaptive ML models, and environmental context awareness in uncontrolled conditions.

Proposed models

This paper proposes a novel EPWOD-POAADP method. The main aim of the technique is to enhance the pedestrian walkway method for blind people's navigation. Figure 1 represents the entire flow of the EPWOD-POAADP model.

Stage I: image pre-processing

At first, the image pre-processing stage applies MF to eliminate the noise in the input data¹⁸. This model is chosen for its robust capability to remove impulsive noise while preserving essential edges and details, which is crucial for accurate object detection and classification. Unlike mean filtering, which can blur edges, MF maintains sharp boundaries, improving the quality of input images. Its nonlinear nature makes it particularly effectual against salt-and-pepper noise commonly found in real-world images. Furthermore, MF is computationally efficient and simple to implement, making it appropriate for real-time applications. These merits collectively justify its selection over other smoothing techniques, ensuring improved downstream model performance.

MF is a nonlinear image processing model frequently employed to decrease noise while maintaining limits in images. In assessing pedestrian walkways to help disabled persons, MF aids in improving the quality of input imageries by eliminating unwanted noise, like distortions from low-quality camera sensors or climate conditions. This pre-processing stage confirms that object detection systems can more precisely classify walkway cracks, obstacles, or other problems. By enhancing the clarity of image data, MF helps measure the availability

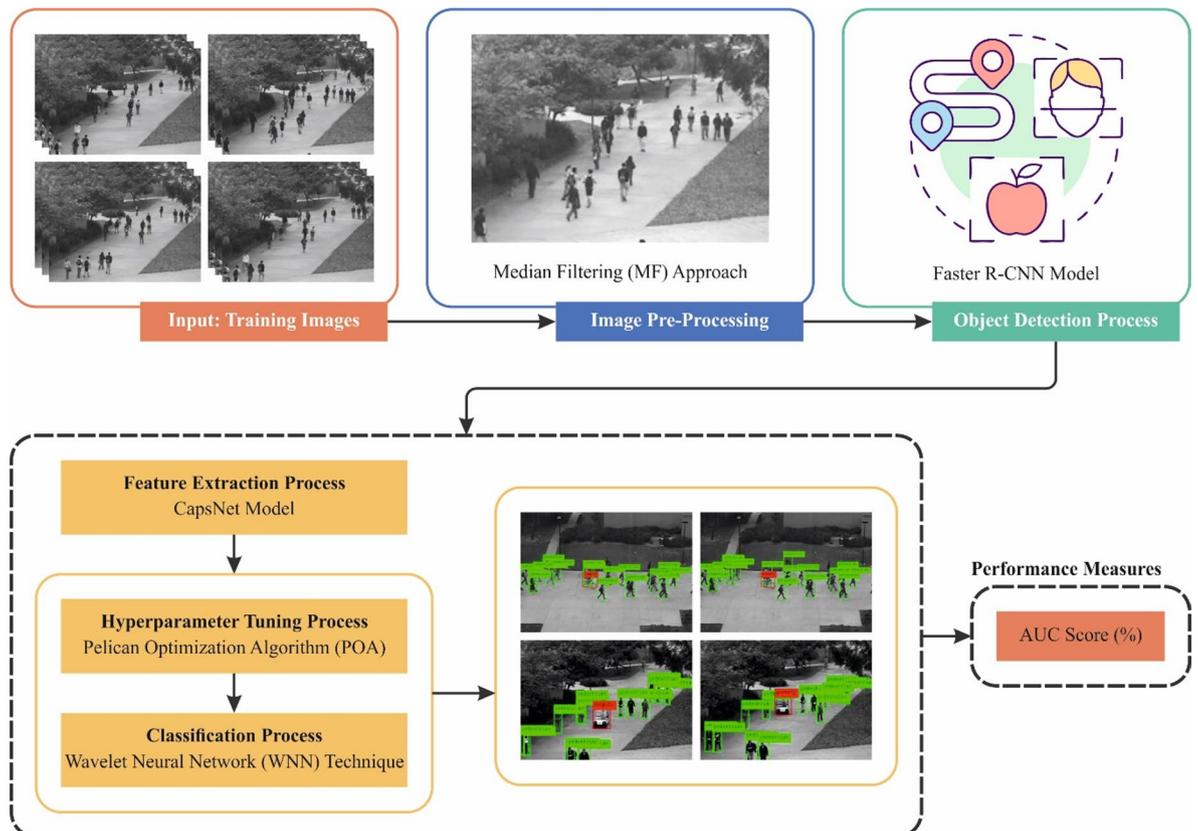


Fig. 1. Overall flow of EPWOD-POAADP model.

and protection of pedestrian tracks for persons with disabilities, eventually donating to better urban planning and substructure development.

Stage II: object detection

Besides, the Faster R-CNN model is employed for the object detection process to identify and locate objects within an image¹⁹. This model is chosen for its excellent balance between accuracy and speed, making it highly appropriate for real-time applications. This model integrates a region proposal network (RPN) that efficiently produces high-quality region proposals, mitigating computational overhead. This end-to-end architecture allows for joint optimization, improving detection precision. Compared to single-stage detectors such as YOLO or SSD, Faster R-CNN generally attains higher accuracy, particularly for detecting small or overlapping objects. Its robustness in handling intrinsic scenes and varying object scales makes it an ideal choice for precise and reliable detection tasks.

Deeper ConvNets are frequently applied for object detection due to their high precision compared to preceding techniques, namely ResNets, VGGNets, DenseNet, and Inception networks. One famous framework is RCNN, which uses deeper ConvNets to identify object applications (possible regions of interest). Though it attains higher precision, it contracts space and time inadequacies. The method captures longer times and needs a larger storage area as it removes characteristics from all images and preserves them on hard disks. The detection procedure only captures 47 s for one image. Faster RCNN considerably enhances the detection speed to 0.3s per image by combining a pooling layer of ROI.

The disadvantage of Fast RCNN is tackled by Faster RCNN, which presents the RPN. This RPN is executed as a complete convolution system, which forecasts object limitations and objectless scores. It attains translation invariance by fastening it with dissimilar ratios and scales. By combining the deeper VGG-16 method, the whole method can effectively carry out the detection and proposal procedure in just 0.2s. This paper recommends an ensemble learning model derived from DL methods for detecting distract drivers. The model attains higher precision by adjusting the Faster RCNN method and removing pose facts from the driver's posture (97.7% validation precision). The method concentrates on objects straightforwardly related to distraction and computes communicating relations utilizing the connection above union metric. It attains a precision of 92.2%, exceeding R-CNN and Faster RCNN. To safeguard its expediency, the paper must assess the real-world performance of the model, reflecting response time and computational efficacy. Another study references an enhanced Faster R-CNN method for smaller object detection. The model presents new methods for ROI pooling and bounding box regression to deal with positioning deviation problems. This specifies the efficiency of Faster RCNN for smaller object detection. Nevertheless, added investigation is essential to assess its performance on dissimilar domains and objects, considering computational complexity and possible drawbacks.

Stage III: feature extraction

For the feature extraction process, the EPWOD-POAADP model employs CapsNet²⁰. This model is chosen as it effectually preserves spatial hierarchies and part-to-whole relationships in visual data, which conventional CNNs often overlook. Unlike standard CNNs, CapsNet utilizes dynamic routing to maintain orientation and pose data, improving robustness to image transformations and distortions. This results in an enhanced generalization, particularly in intrinsic scenarios where the spatial arrangement of features is crucial. Moreover, CapsNet requires fewer training samples to achieve high accuracy, making it effective in data-scarce environments. Its capability to capture richer feature representations presents a significant advantage over conventional feature extractors. Figure 2 exemplifies the structure of CapsNet.

An NN named CapsNet was recently presented, and it could considerably influence DL, mainly in computer vision (CV). The output and input of the neuron in a traditional CNN are scalars. On the other hand, the vector is handled by the neurons in CapsNet. Therefore, the capsule is otherwise called a vector neuron (VN), and a vector encompasses all essential data concerning the status of the features in the capsule recognition method. After resizing and deleting features, pooling layers of CNN drop numerous essential features. Furthermore, a CNN fails to understand relationships amongst numerous removed features due to the function, which might obtain crucial data that does not appear. CapsNet utilizes squash functions in association with pooling layers. Like a nonlinear function, which captures input using the vector model and resizes data in the unit vector without changing its alignment, this task will not cause some data to get lost. The following encloses a calculated equation for the capsule's operation,

$$\hat{u}_{j|i} = W_{ij}u_i \quad (1)$$

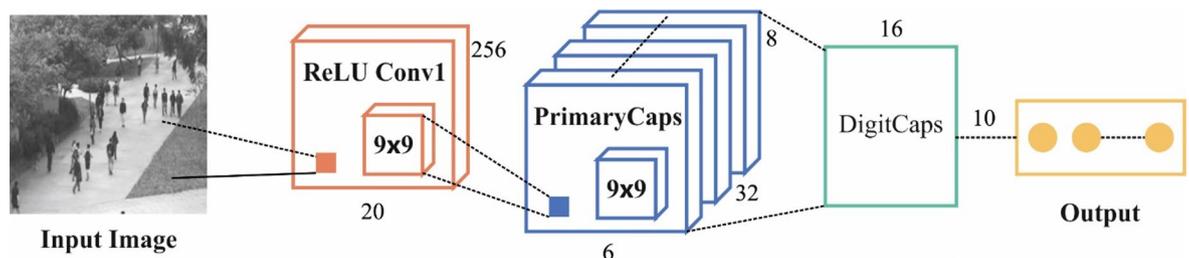


Fig. 2. Structure of CapsNet.

The prediction vector is represented as $\hat{u}_{j|i}$, acknowledged by capsule j and produced by capsule i . This multiplies the weighted matrix W_{ij} by the output u_i of the capsule layer that came before it.

$$s_j = \sum_i c_{ij} \hat{u}_{j|i} \quad (2)$$

The total product counts $\hat{u}_{j|i}$ and c_{ij} give outcomes in s_j . During CapsNet, capsules were applied instead of conventional CNN neurons, and all input and output units were transformed into vectors. The vector's orientation designates a specific unit's influences on the input data. The capsule vector size designates an object's possible existence in the present input. The activation function of the CNN, or another *squashing* function, guarantees that the vector length is amongst (0,1). Equation (3) can define the *squashing* function.

$$v_j = \frac{\|S\|}{1 + \|S\|^2} \frac{S_j}{\|S_j\|} \quad (3)$$

Simultaneously, the capsule's complete input vector is represented as S_j and the capsule j 's output vector is shown as v_j .

$$c_{ij} = \frac{\exp(b_{ij})}{\sum_j \exp(b_{ij})} \quad (4)$$

The dynamical routing method describes the coupling coefficient C in Eq. (10). The softmax function is defined as c_{ij} and. It specifies the log prior prospect amongst capsules i and j . In previous layers, CapsNet applied the parameter b_{ij} to identify relations amongst capsules i and j . The coupling coefficient C is equivalent to complete capsules in a layer, and an initial iteration b_{ij} is set to 0. Equation (11) was applied to update \hat{u} and v_j . Utilizing the dot product of v_j and \hat{u} , the following equation updates the parameter b_{ij} :

$$b_{ij} = b_{ij} + \hat{u}_{j|i} v_j \quad (5)$$

The b_{ij} value will improve after updating utilizing Eq. (11) after the $\hat{u}_{j|i}$ and v_j dot product gives an optimistic outcome. By strengthening the bond among capsules i and j , greater b_{ij} leads to better, making greater S_j and v_j values. There should be harm to the connection between capsules i and j when the dot product of v_j and \hat{u} is negative.

Stage IV: pedestrian walkway detection using WNN

Furthermore, the WNN technique is implemented for the detection and classification²¹. This technique was chosen for its robust capability in capturing both time-frequency information and nonlinear relationships within data, which is significant for handling complex patterns in pedestrian environments. This model integrates wavelet transform, effectively analyzing localized features and image variations. This results in an enhanced accuracy in detecting walkways, particularly in noisy or cluttered scenes. Moreover, WNN demonstrates faster convergence and better generalization with fewer parameters, making it computationally efficient and appropriate for real-time applications. Its ability to balance precision and speed presents a clear advantage over other detection models.

A WNN establishes higher learning abilities, quicker convergence rates, and better accuracy than conventional BP neural networks and other feed-forward neural networks. Together with enhanced sensitivity in function calculation and strong fault tolerance, these benefits make WNNs mainly efficient in tackling composite signal denoising tasks. For this paper, the powers of WNNs is utilized by incorporating wavelet transforms' multiple-scale study with NN's nonlinear capacity to handle. This hybrid method permits WNNs to adaptively take signal dissimilarities through dissimilar scales, allowing effective processing of either higher- or lower-frequency elements. During this figure, X_1, X_2, \dots, X_m characterize the input parameters of the WNN, whereas Y_1, Y_2, \dots, Y_l represent the forecast output values. ω_{jk} and ω_{ij} indicate the corresponding connection weighting between the input and hidden layers (HL) and between the HL and the output layers.

$$h(j) = h_j \left(\frac{\sum_{i=1}^n \omega_{ij} x_i - b_j}{a_i} \right) j = 1, 2, \dots, l \quad (6)$$

If the sequence of the input signal is $X_i (i = 1, 2, \dots, /k)$, the output equation for the HL is as demonstrated:

In Eq. (6): $h(j)$ refers to an output value of the j node in the HL. h_j stands for wavelet basis function. a_i , and b_j represents scaling and translation factor s. The computation equation is demonstrated below:

$$y(k) = \sum_{i=1}^l \omega_{jk} h(i) k = 1, 2, \dots, m \quad (7)$$

In Eq. (7): $h(i)$ denotes the output value of the i HL. l means HL node counts. m refers to output layer node counts. The WNN typically utilizes the gradient correction model to correct the network weighting and wavelet base function parameters. The correction method is shown below:

Compute the prediction error of WNN:

$$e = \sum_{k=1}^m yn(k) - y(k) \quad (8)$$

In Eq. (8), $yn(k)$ is the predictable output, and $y(k)$ refers to the projected output of the WNN. Correct the weighting of WNN based on the prediction error:

$$\omega_{n,k}^{(i+1)} = \omega_{n,k}^i + \Delta \omega_{n,k}^{(i+1)} \quad (9)$$

The coefficients of wavelet base functions are modified based on the prediction error e :

$$a_k^{(i+1)} = a_k^i + \Delta a_k^{(i+1)} \quad (10)$$

$$b_k^{(i+1)} = b_k^i + \Delta b_k^{(i+1)} \quad (11)$$

In Eq. (9)–(11), $\Delta \omega_{n,k}^{(i+1)}$, $\Delta a_k^{(i+1)}$, and $\Delta b_k^{(i+1)}$ are computed by the networking prediction error. The computation model is as demonstrated:

$$\Delta \omega_{n,k}^{(i+1)} = -\eta \frac{\partial e}{\partial \omega_{n,k}^{(i)}} \quad (12)$$

$$\Delta a_k^{(i+1)} = -\eta \frac{\partial e}{\partial a_k^{(i)}} \quad (13)$$

$$\Delta b_k^{(i+1)} = -\eta \frac{\partial e}{\partial b_k^{(i)}} \quad (14)$$

Whereas η refers to the networking rate of learning.

Stage V: POA-based parameter tuning

Finally, the hyperparameter range of the WNN model is performed by implementing the POA method²². This model is chosen for its excellent balance between exploration and exploitation capabilities, effectively searching the hyperparameter space for optimal values. Compared to conventional optimization methods and other metaheuristics, POA illustrates faster convergence and avoids getting trapped in local minima, resulting in improved overall model performance. Its simple yet efficient mechanism allows it to handle complex, multi-dimensional problems with fewer computational resources. Additionally, POA's adaptability and robustness make it appropriate for tuning parameters in DL models like WNN, ensuring improved accuracy and stability without excessive computational overhead.

All population members specify candidate solutions, and the optimization problem variables are based on their location inside the space. At the starting phase, Eq. (15) specified population members at the upper and lower limits of the problem.

$$x_{i,j} = l_j + rand \cdot (u_j - l_j), \quad i = 1, 2, \dots, N, j = 1, 2, \dots, m \quad (15)$$

Whereas $x_{i,j}$ refers to the value of the j_{th} variable identified by the i_{th} candidate solution, N stands for population member count, m denotes problem variable amount, $rand$ signifies the number generated at random in the interval (0,1), l_j and u_j represent the j_{th} lower and upper limit of problem variables. The hunting tactic is modelled in dual phases, such as the exploration and exploitation stages.

In the exploration stage, the pelicans find the prey and approach it. This theory is mathematically pretended in Eq. (16).

$$x_{i,j}^{p1} = \begin{cases} x_{i,j} + rand \cdot (p_j - I \chi_{i,j}), & F_p < F_i; \\ x_{i,j} + rand \cdot (x_{i,j} - p_j), & else \end{cases} \quad (16)$$

Whereas $x_{i,j}^{p1}$ refers to the novel status of the i_{th} pelican in the j_{th} size according to stage 1, p_j stands for the position of prey in the j_{th} size, and F_p is its value of the objective function. I mean a number that is randomly equivalent to 1 or 2 and arbitrarily chosen for all iterations and all members.

In the exploitation stage, once the pelicans reach the water's surface, they spread their wings and travel near the fish to a shallow region for collection. The pelican's behaviour in searching is pretended mathematically in Eq. (17).

$$x_{i,j}^{p2} = x_{i,j} + R \cdot \left(1 - \frac{t}{T}\right) \cdot (2 \cdot rand - 1) x_{i,j} \quad (17)$$

Whereas $x_{i,j}^{p1}$ stands for the present status of the i_{th} pelican in the j_{th} size according to stage 2, R denotes constant equivalent to 0.2, $R \cdot \left(1 - \frac{t}{T}\right)$ epitomizes the neighbourhood radius of $x_{i,j}$, t represents the iteration counter, and T symbolizes maximal iteration counts.

Dataset	Videos	Frames of Average	Length
“UCSDPed1 (Bikers, small carts, walking across walkways)”	70	201	5 min
“UCSDPed2 (Bikers, small carts, walking across walkways)”	28	163	5 min

Table 1. Details on the dataset.

TPR					
FPR	MPPCA	Social Force	EADN	ADPW-FLHHO	EPWOD-POAADP
0	0.0000	0.0000	0.0000	0.0000	0.0000
5	0.0915	0.1315	0.3466	0.5958	0.7129
10	0.2269	0.2455	0.5506	0.7445	0.8606
15	0.3517	0.3676	0.7547	0.8239	0.8906
20	0.4311	0.4578	0.7550	0.8744	0.9017
25	0.9379	0.9218	0.5373	0.5188	0.9523
30	0.5746	0.6329	0.9431	0.9619	0.9719
35	0.6353	0.7175	0.9565	0.9725	0.9865
40	0.6805	0.8104	0.9646	0.9775	0.9805
45	0.7549	0.8821	0.9726	0.9831	0.9904
50	0.7972	0.9089	0.9776	0.9857	1.0000
55	0.8319	0.9324	0.9776	0.9856	1.0000
60	0.8796	0.9409	0.9778	0.9882	1.0000
65	0.9115	0.9486	0.9831	0.9881	1.0000
70	0.9540	0.9566	0.9829	1.0000	1.0000
75	0.9566	0.9751	0.9885	1.0000	1.0000
80	0.9616	0.9831	1.0000	1.0000	1.0000
85	0.9671	0.9829	1.0000	1.0000	1.0000
90	0.9833	0.9831	1.0000	1.0000	1.0000
95	0.9805	0.9938	1.0000	1.0000	1.0000
100	1.0000	1.0000	1.0000	1.0000	1.0000

Table 2. Comparative analysis of EPWOD-POAADP technique with other approaches below UCSDPed1 dataset.

Therefore, POA meets solutions quicker to the global optimum-based and successfully upgrades to reject or accept the novel pelican location. The POA originates from a fitness function (FF) for attaining an enhanced classification performance. It expresses a positive numeral to epitomize the better result of the candidate solution. The classification rate of error reduction was measured as FF. Its mathematical formulation is computed in Eq. (18).

$$\begin{aligned}
 \text{fitness}(x_i) &= \text{ClassifierErrorRate}(x_i) \\
 &= \frac{\text{no. of misclassified samples}}{\text{Total no. of samples}} \times 100
 \end{aligned}
 \quad (18)$$

Performance analysis

The performance evaluation of the EPWOD-POAADP methodology is examined using the UCSD anomaly detection dataset. The technique is simulated using Python 3.6.5 on a PC with an i5-8600k, 250GB SSD, GeForce 1050Ti 4GB, 16GB RAM, and 1 TB HDD. Parameters include a learning rate of 0.01, ReLU activation, 50 epochs, 0.5 dropouts, and a batch size of 5. Table 1 represents a detailed description of the dataset.

Table 2; Fig. 3 show the overall comparative results of the EPWOD-POAADP approach with existing methods under the UCSDPed1 dataset²³. The table values implied that the EPWOD-POAADP approach exhibited effective performances. Based on five false positive rates (FPR), the EPWOD-POAADP model has obtained a higher true positive rate (TPR) of 0.7129 while the MPPCA, SF, EADN, and ADPW-FLHHO models achieved lesser TPR of 0.0915, 0.1315, 0.3466, and 0.5958. Followed by, depending on 15 FPR, the EPWOD-POAADP technique gained a better TPR of 0.8906 whereas the MPPCA, SF, EADN, and ADPW-FLHHO models attained a lower TPR of 0.3517, 0.3676, 0.7547, and 0.8239. In addition, for 25 FPR, the EPWOD-POAADP approach has achieved a greater TPR of 0.9523 whereas the MPPCA, SF, EADN, and ADPW-FLHHO models have gained the worst TPR of 0.9379, 0.9218, 0.5373, and 0.5188. Moreover, based on 50 FPR, the EPWOD-POAADP approach has gotten a superior TPR of 1.0000 while the MPPCA, SF, EADN, and ADPW-FLHHO models accomplished an inferior TPR of 0.7972, 0.9089, 0.9776, and 0.9857. Finally, depending on 60 FPR, the EPWOD-POAADP

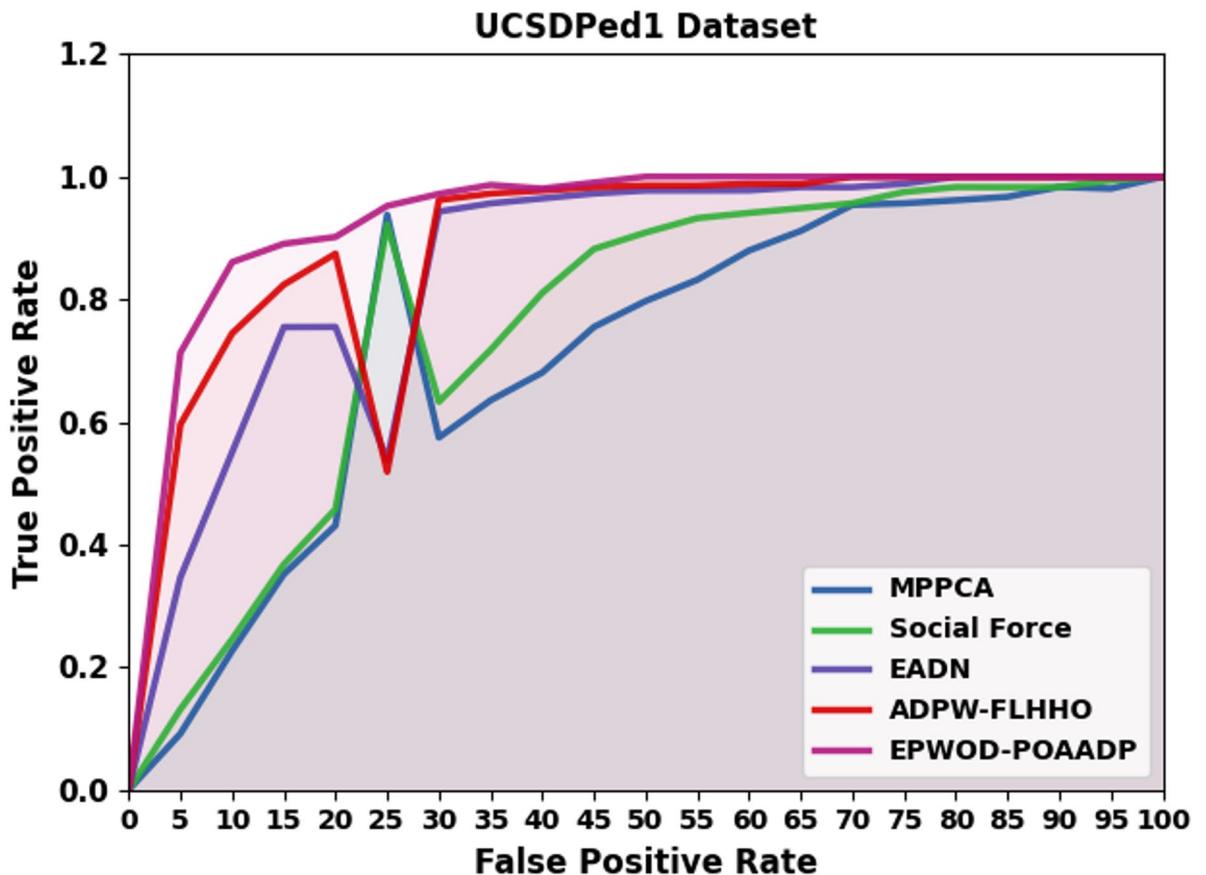


Fig. 3. Comparative outcome of EPWOD-POAADP technique under UCSDPed1 dataset.

method has achieved a maximal TPR of 1.0000 whereas the MPPCA, SF, EADN, and ADPW-FLHHO models attained a lower TPR of 0.8796, 0.9409, 0.9778, and 0.9882.

Figure 4 illustrates the TRA $accu_y$ (TRAAY) and validation $accu_y$ (VLAAY) analysis of the EPWOD-POAADP technique below the UCSDPed1 dataset. The $accu_y$ analysis is calculated across an interval of 0–50 epochs. The figure highlights that the TRAAY and VLAAY values exhibit an increasing trend, which informs the capacity of the EPWOD-POAADP technique, which has superior performance across multiple iterations. In addition, the TRAAY and VLAAY leftovers are closer across the epochs, which specifies inferior overfitting and displays the maximum performance of the EPWOD-POAADP technique, guaranteeing reliable prediction on hidden samples.

In Fig. 5, the EPWOD-POAADP methodology's TRA loss (TRALO) and VLA loss (VLALO) display under the UCSDPed1 dataset is demonstrated. The loss values are computed over the range of 0–50 epochs. The TRALO and VLALO values are intended to exemplify a diminishing trend, which informs the method's capability in balancing a trade-off.

Table 3; Fig. 6 report a detailed AUC_{score} study of the EPWOD-POAADP technique below the UCSDPed1 dataset²⁴. The outcomes illustrated that the TSN-RGB, Spatiotemporal, and TSN-Optical Flow techniques have displayed ineffectual outcomes with the least AUC_{score} of 90.57%, 91.64%, and 92.91%, individually. In the meantime, the MIL-C3D, Binary SVM, and EADN techniques have shown significant performance with AUC_{score} of 95.05%, 96.78%, and 98.41%. Likewise, the ADPW-FLHHO techniques have accomplished reasonable results with AUC_{score} of 99.40%. Besides, the EPWOD-POAADP method proves higher performance with a better AUC_{score} of 99.51%.

Table 4; Fig. 7 illustrate the computational time (CT) analysis of the EPWOD-POAADP technique with existing models under the UCSDPed1 dataset. The EPWOD-POAADP technique illustrates the most efficient performance with a CT of 6.39 s, exhibiting a significant improvement over other model. For instance, the ADPW-FLHHO and Binary SVM Method register CTs of 9.63 and 10.45 s, respectively, while the MIL-C3D model and EADN Method exhibit higher CTs of 11.36 and 13.42 s. Furthermore, the TSN-Optical Flow method records 12.91 s, the Spatiotemporal model 12.41 s, and the TSN-RGB 8.41 s. The EPWOD-POAADP model's reduced CT highlights its suitability for time-sensitive applications, presenting faster processing without compromising performance AUC_{score} of 99.03%.

Table 5; Fig. 8 describe the ablation study of the EPWOD-POAADP approach with the existing models under the UCSDPed1 dataset. The EPWOD-POAADP approach achieved the highest AUC_{score} of 99.51%, significantly outperforming the existing models such as WNN with 98.62%, POA with 98.10%, and CapsNet

Training and Validation Accuracy - UCSDPed1 Dataset

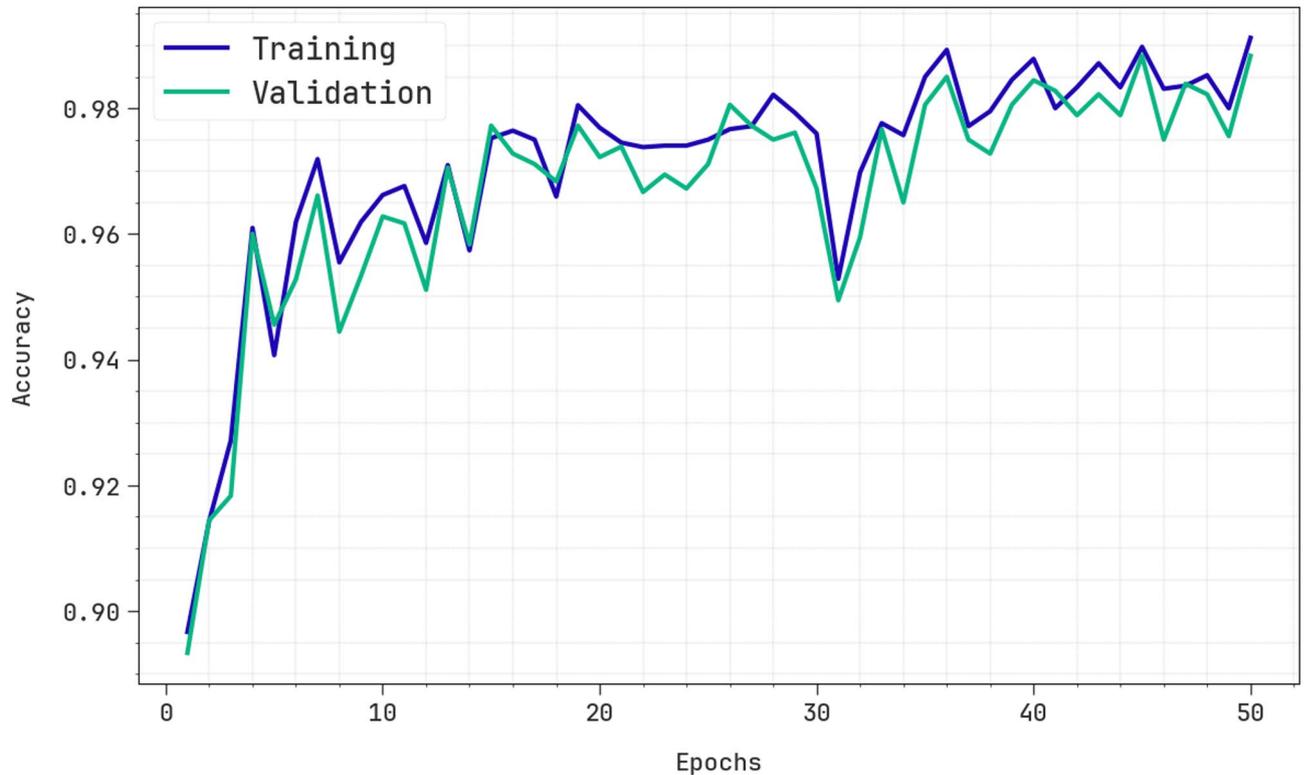


Fig. 4. $Accu_y$ outcome of EPWOD-POAADP technique under UCSDPed1 dataset.

with 97.34%. Conventional approaches like Faster R-CNN and MF attained lesser with AUC_{score} of 96.80% and 96.00%, subsequently. These outputs emphasize the superior anomaly detection capability of the EPWOD-POAADP model and confirm that its enhancements contribute meaningfully to performance gains over both classical and DL-based methods.

Table 6; Fig. 9 show the overall comparative outcomes of the EPWOD-POAADP technique with the existing methods below the UCSDPed2 dataset. The table values suggest that the EPWOD-POAADP technique showed the effectual performances. Depending on 5 FPR, the EPWOD-POAADP technique has gained a greater TPR of 0.7410 whereas the MPPCA, SF, EADN, and ADPW-FLHHO methods have attained lower TPR of 0.0761, 0.1287, 0.3483, and 0.5720. Followed, concerning 15 FPR, the EPWOD-POAADP technique has achieved a greater TPR of 0.9497, whereas the MPPCA, SF, EADN, and ADPW-FLHHO methods have accomplished a minimal TPR of 0.3660, 0.4204, 0.6053, and 0.7937. In addition, depending on 25 FPR, the EPWOD-POAADP technique attained a better TPR of 0.9253 while the MPPCA, SF, EADN, and ADPW-FLHHO approaches realized the worst TPR of 0.5624, 0.6508, 0.7838, and 0.9095. Additionally, for 50 FPR, the EPWOD-POAADP method has gained a greater TPR of 1.0000 whereas the MPPCA, SF, EADN, and ADPW-FLHHO approaches have reached inferior TPR of 0.8333, 0.9455, 0.9692, and 0.9858. Lastly, based on 55 FPR, the EPWOD-POAADP method has gained a superior TPR of 1.0000 while the MPPCA, SF, EADN, and ADPW-FLHHO approaches achieved the worst TPR of 0.9283, 0.9575, 0.9820, and 0.9914.

Figure 10 illustrates the TRAAy and VLAAY analysis of the EPWOD-POAADP technique below the UCSDPed2 dataset. The $accu_y$ values are computed within the range of 0–50 epochs. The figure highlights that the TRAAy and VLAAY analysis exhibits an increasing trend, which informed the capacity of the EPWOD-POAADP methodology with maximum performance across several iterations. Simultaneously, the TRAAy and VLAAY remain closer across the epochs, identifying inferior overfitting and exhibiting greater performance of the EPWOD-POAADP technique, promising reliable prediction on hidden samples.

Figure 11 illustrates the TRALO and VLALO curves of the EPWOD-POAADP approach under the UCSDPed2 dataset is displayed. The loss values are computed within the range of 0–50 epochs. It signifies that the TRALO and VLALO values establish a reducing trend, which informs the capacity of the EPWOD-POAADP method to balance a trade-off.

In Table 7; Fig. 12, a thorough AUC_{score} experiment of the EPWOD-POAADP methodology below the UCSDPed2 dataset is reported correctly. The outcomes illustrated that the TSN-RGB, Spatiotemporal, and TSN-Optical Flow techniques have displayed ineffectual outcomes with lower AUC_{score} of 90.45%, 92.49%, and 94.37%, respectively. In the meantime, the MIL-C3D, Binary SVM, and EADN techniques have demonstrated large performance with AUC_{score} of 95.51%, 97.17%, and 98.31%. Furthermore, the ADPW-FLHHO approach

Training and Validation Loss - UCSDPed1 Dataset

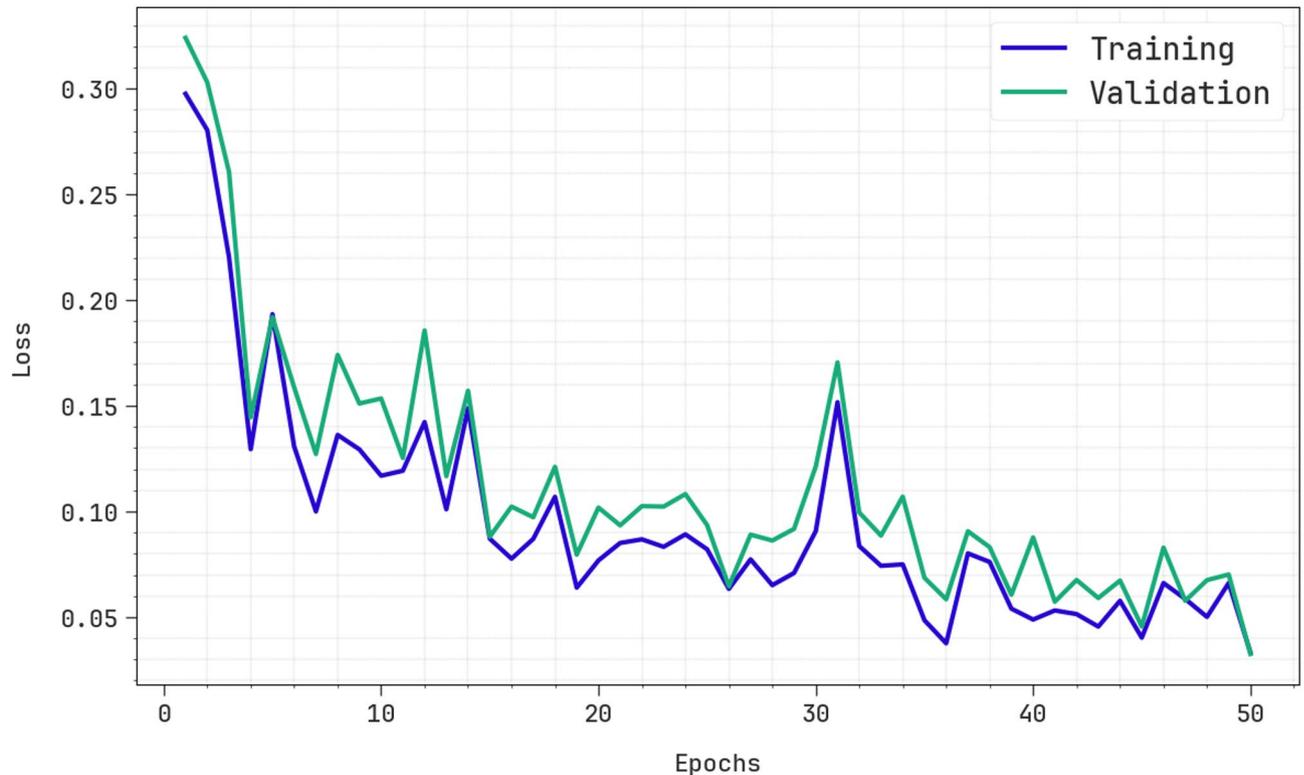


Fig. 5. Loss outcome of EPWOD-POAADP technique under UCSDPed1 dataset.

Models	AUC_{score}
EPWOD-POAADP	99.51
ADPW-FLHHO	99.40
EADN method	98.41
Binary SVM method	96.78
MIL-C3D model	95.05
TSN-optical flow method	92.91
Spatiotemporal model	91.64
TSN-RGB	90.57

Table 3. AUC_{score} Outcome of EPWOD-POAADP method with existing models under UCSDPed1 dataset.

has accomplished reasonable results with AUC_{score} of 99.20%. Finally, the EPWOD-POAADP approach exhibited maximum performance with an increased AUC_{score} of 99.35%.

Table 8; Fig. 13 specify the CT analysis of the EPWOD-POAADP methodology with the existing models under the UCSDPed2 dataset. The EPWOD-POAADP methodology achieves a CT of 8.12 s, outperforming all comparative approaches and highlighting its optimized execution speed. In contrast, the ADPW-FLHHO, EADN, and Binary SVM models report slower CTs of 11.34, 11.23, and 11.76 s respectively. The MIL-C3D and TSN-RGB approach exhibit CTs of 12.47 and 11.87 s, while the TSN-Optical Flow system and Spatiotemporal method are considerably slower with CTs of 13.72 and 19.52 s. The reduced CT of the EPWOD-POAADP method assists its suitability for latency-critical applications, presenting fast decision-making with a high AUC_{score} of 99.03%. This rapid responsiveness makes it ideal for real-time pedestrian safety systems, especially in dynamic urban environments.

Table 9; Fig. 14 depict the ablation study of the EPWOD-POAADP methodology with the existing models under the UCSDPed2 dataset. The EPWOD-POAADP methodology attained an AUC_{score} of 99.35%, clearly outperforming the existing techniques such as WNN with 98.46%, POA with 97.71%, and CapsNet with 96.95%. Meanwhile, Faster R-CNN and MF achieved lesser AUC_{score} of 96.45% and 95.81%, correspondingly. These outputs demonstrate that the EPWOD-POAADP model provides superior anomaly detection performance,

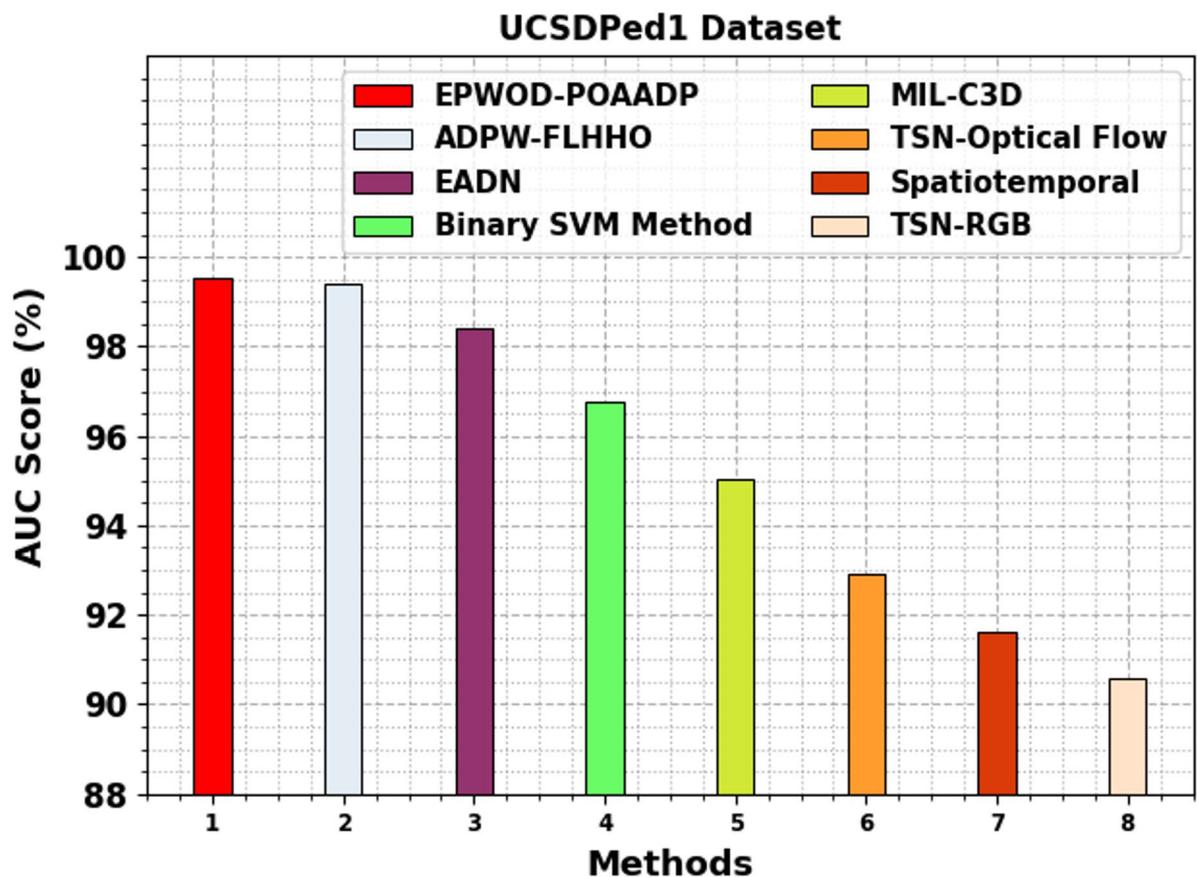


Fig. 6. AUC_{score} analysis of EPWOD-POAADP method under UCSDPed1 dataset.

Models	CT (sec)
EPWOD-POAADP	6.39
ADPW-FLHHO	9.63
EADN Method	13.42
Binary SVM Method	10.45
MIL-C3D model	11.36
TSN-Optical Flow method	12.91
Spatiotemporal model	12.41
TSN-RGB	8.41

Table 4. CT analysis of EPWOD-POAADP technique with existing models under UCSDPed1 dataset.

validating the impact of its architectural innovations and optimization strategy in handling complex video surveillance data.

Table 10 indicates the ablation study comparing the computational efficiency of diverse upsampling methods in terms of FLOPs and GPU memory consumption²⁵. The EPWOD-POAADP method attained the lowest FLOPs at 90.34 and the lowest GPU usage at 1200, significantly outperforming all other methods. In contrast, Pixel Shuffle recorded the highest FLOPs at 167.31, while Dysample consumed the most GPU memory at 3530. Other methods like Deconv and Bilinear illustrated relatively higher resource demands, with FLOPs of 143.93 and 135.86, and GPU usage of 2748 and 3049 respectively. These results emphasize that the EPWOD-POAADP model is not only computationally efficient but also highly appropriate for resource-constrained environments.

Conclusion

In this paper, a novel EPWOD-POAADP method is proposed. The main intention of the EPWOD-POAADP method is to enhance the pedestrian walkways method for blind people's navigation. At first, the image pre-processing stage applies MF to eliminate the noise in the input data. Besides, the Faster R-CNN model is employed for the object detection process to identify and locate objects within an image. The proposed EPWOD-POAADP model designs the CapsNet model to extract the feature method. Furthermore, the WNN technique is

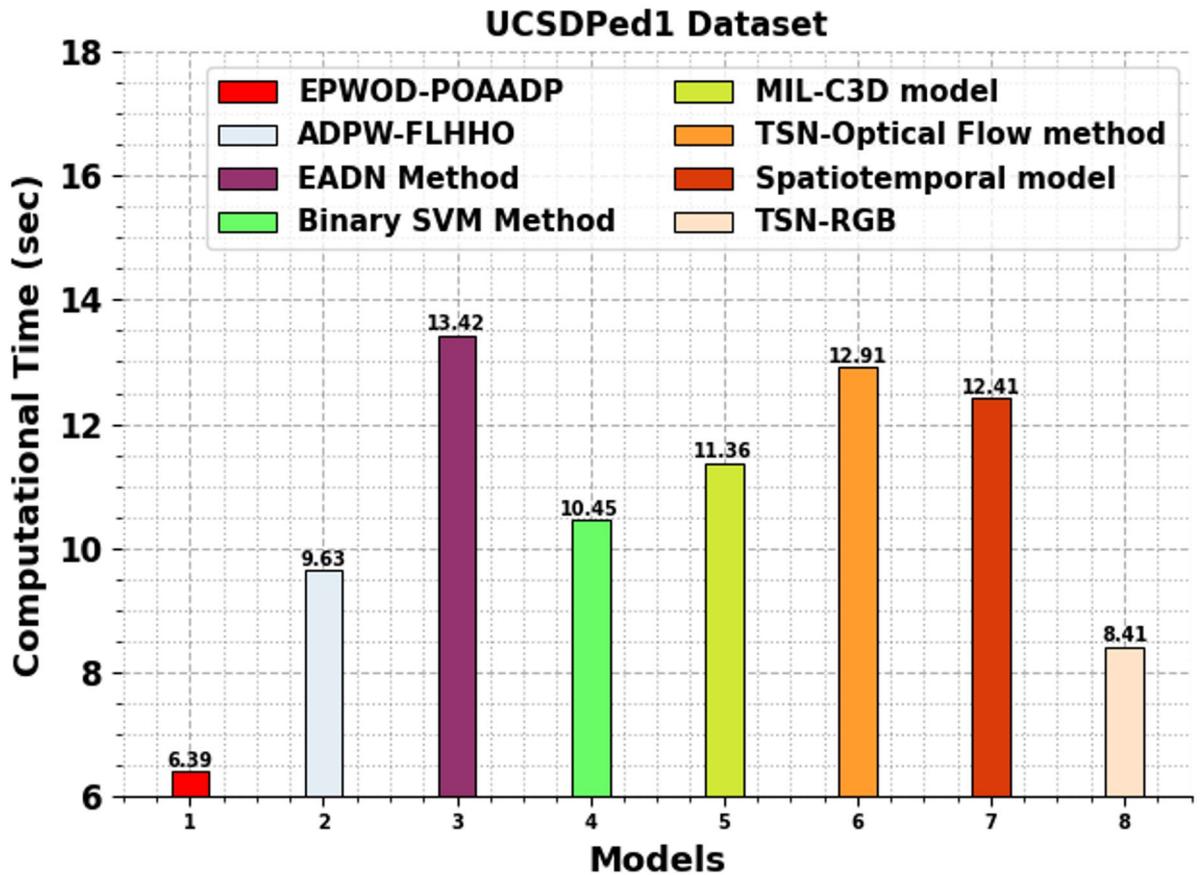


Fig. 7. CT analysis of EPWOD-POAADP technique with existing models under UCSDPed1 dataset.

UCSDPed1 Dataset	
Models	AUC-Score (%)
EPWOD-POAADP	99.51
WNN	98.62
POA	98.10
CapsNet	97.34
Faster R-CNN	96.80
MF	96.00

Table 5. Result analysis of the ablation study of EPWOD-POAADP approach under the UCSDPed1 dataset.

implemented for the detection and classification process. Finally, the POA model performs the hyperparameter range of the WNN model. The experimental evaluation of the EPWOD-POAADP approach is examined using a benchmark image dataset. The results indicated the enhanced performance of the EPWOD-POAADP approach compared to recent approaches. The limitations of the EPWOD-POAADP approach comprise a reliance on a limited dataset, which may affect the generalizability of the results across diverse real-world scenarios. Furthermore, the approach does not address real-time processing constraints, which are significant for practical deployment in dynamic environments. The robustness of the model against varying environmental conditions and occlusions remains unexplored. Furthermore, the scalability of larger and more complex pedestrian networks is not thoroughly evaluated. Future work could explore integrating adaptive learning methods to improve model flexibility, incorporate multi-modal sensor data for improved result, and develop lightweight algorithms suitable for edge computing devices to enable faster, on-site processing.

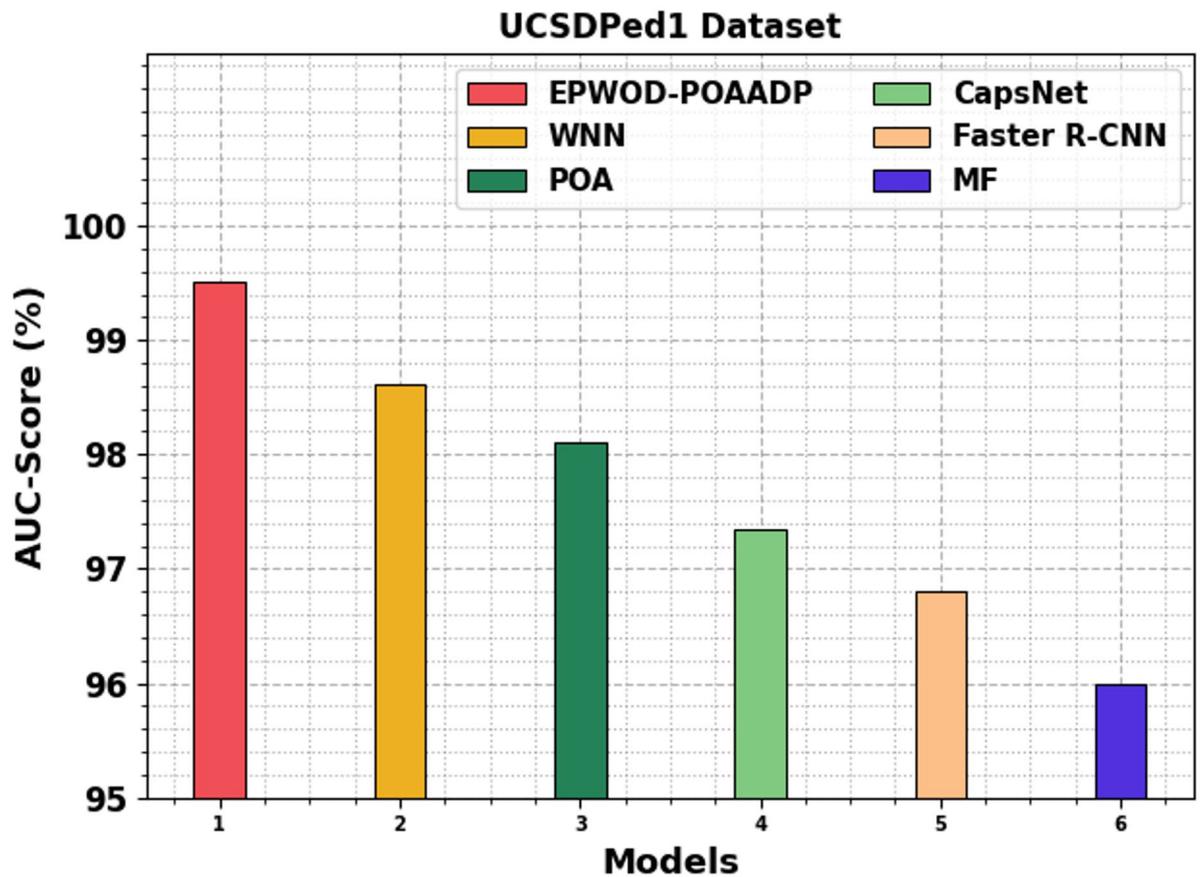


Fig. 8. Result analysis of the ablation study of EPWOD-POAADP approach under the UCSDPed1 dataset.

TPR					
FPR	MPPCA	Social Force	EADN	ADPW-FLHHO	EPWOD-POAADP
0	0.0000	0.0000	0.0000	0.0000	0.0000
5	0.0761	0.1287	0.3483	0.5720	0.7410
10	0.2530	0.2711	0.5350	0.6467	0.8107
15	0.3660	0.4204	0.6053	0.7937	0.9497
20	0.4904	0.4968	0.7487	0.8723	0.9299
25	0.5624	0.6508	0.7838	0.9095	0.9253
30	0.6937	0.7362	0.9277	0.9610	0.9888
35	0.7196	0.8319	0.9500	0.9616	0.9896
40	0.7756	0.8802	0.9557	0.9698	0.9861
45	0.7997	0.9283	0.9630	0.9825	1.0000
50	0.8333	0.9455	0.9692	0.9858	1.0000
55	0.9283	0.9575	0.9820	0.9914	1.0000
60	0.9410	0.9622	0.9867	1.0000	1.0000
65	0.9525	0.9769	0.9964	1.0000	1.0000
70	0.9643	0.9843	0.9912	1.0000	1.0000
75	0.9840	0.9906	0.9904	1.0000	1.0000
80	0.9908	1.0000	1.0000	1.0000	1.0000
85	0.9954	1.0000	1.0000	1.0000	1.0000
90	1.0000	1.0000	1.0000	1.0000	1.0000
95	1.0000	1.0000	1.0000	1.0000	1.0000
100	1.0000	1.0000	1.0000	1.0000	1.0000

Table 6. Comparative result of EPWOD-POAADP technique with other methods under the UCSDPed2 dataset.

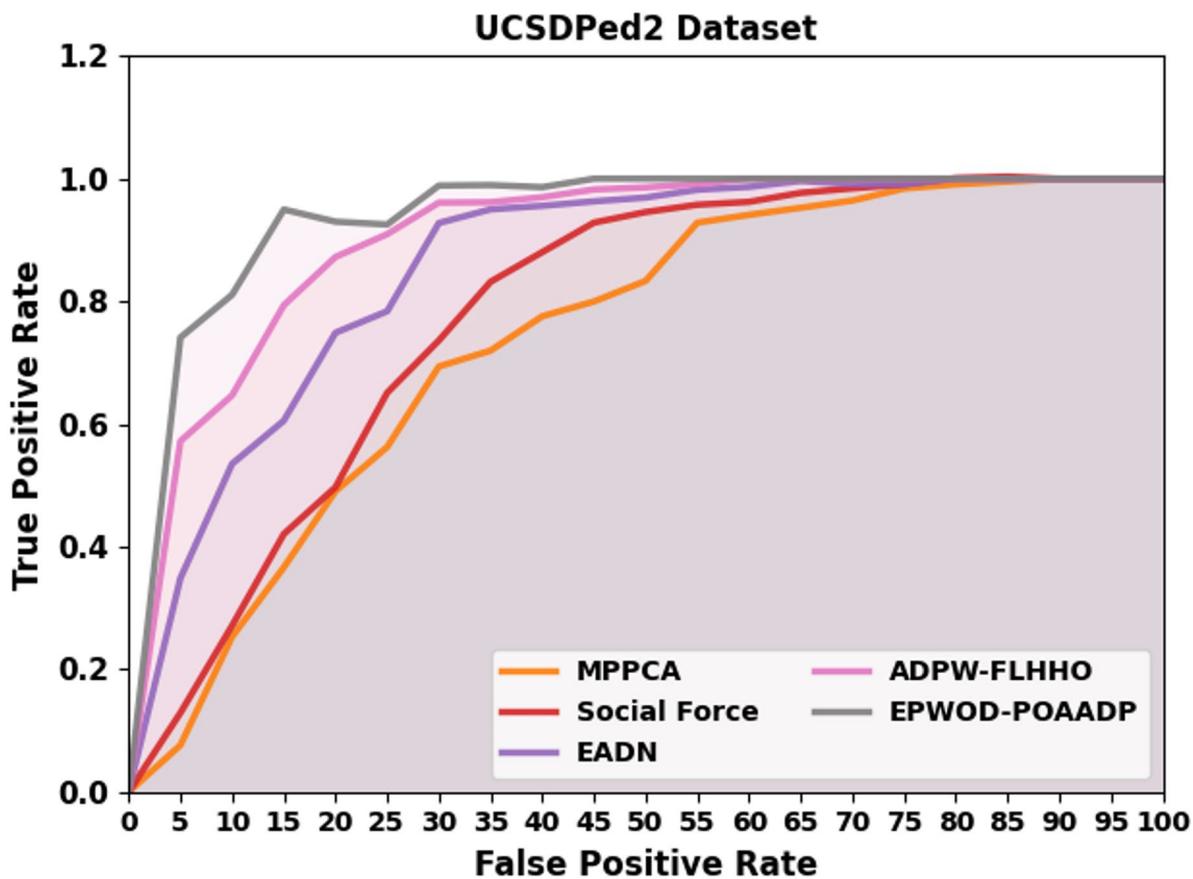


Fig. 9. Comparative outcome of EPWOD-POAADP technique under UCSDPed2 dataset.

Training and Validation Accuracy - UCSDPed2 Dataset

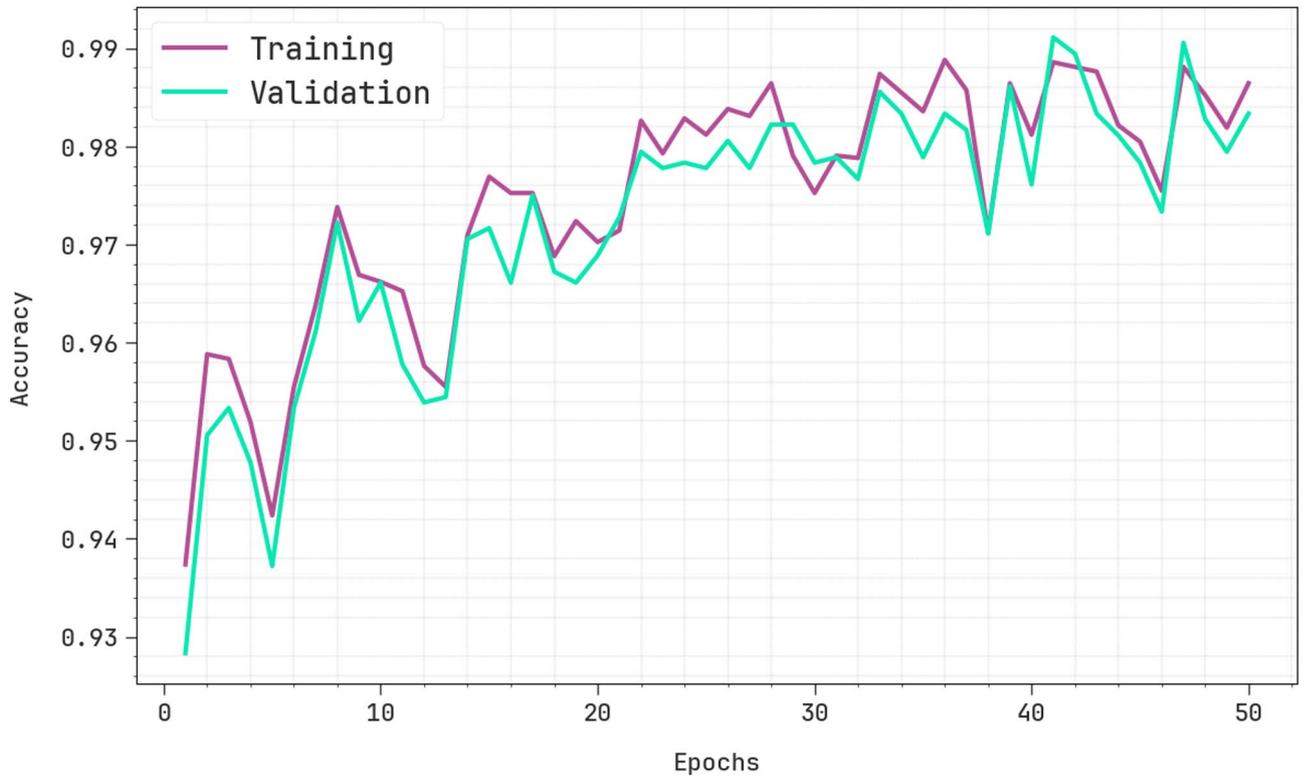


Fig. 10. Accuracy curve of EPWOD-POAADP technique under UCSDPed2 dataset.

Training and Validation Loss - UCSDPed2 Dataset

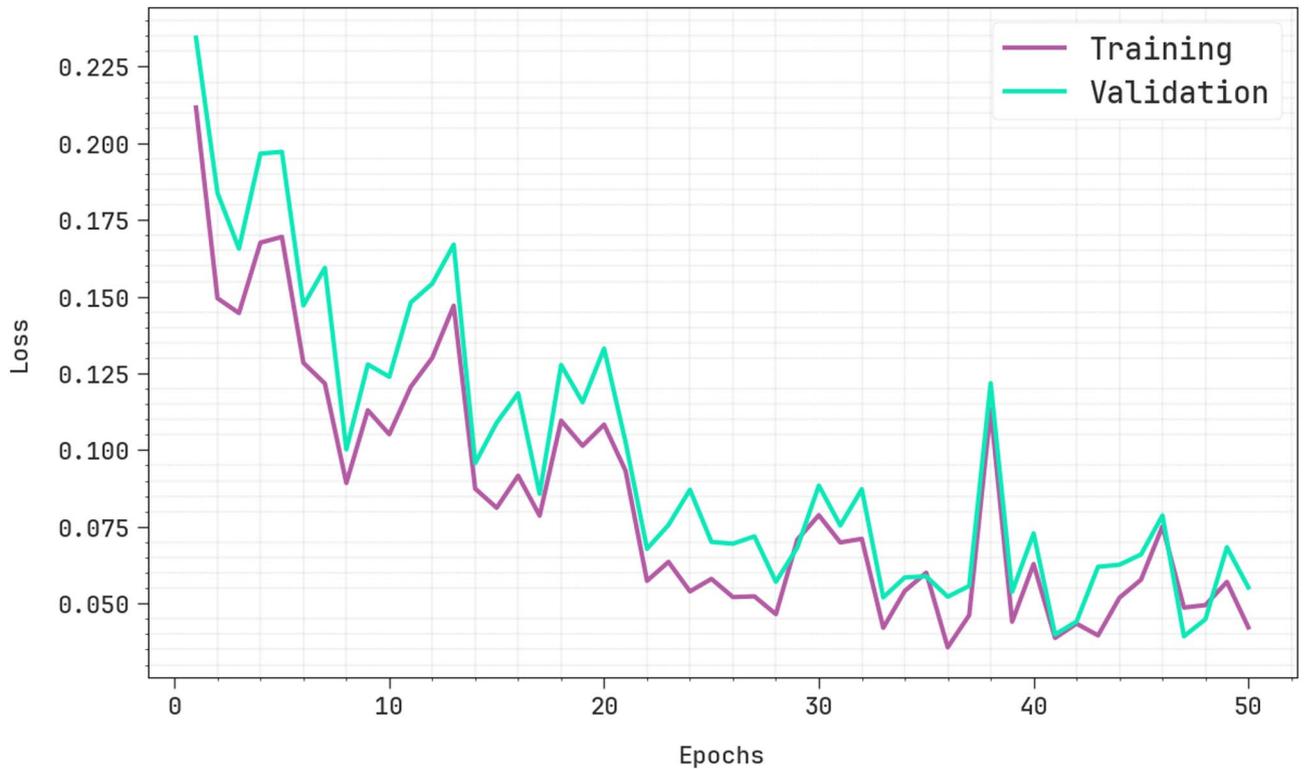


Fig. 11. Loss analysis of EPWOD-POAADP technique below UCSDPed2 dataset.

Methods	AUC Score (%)
EPWOD-POAADP	99.35
ADPW-FLHHO model	99.20
EADN method	98.31
Binary SVM Method	97.17
MIL-C3D technique	95.51
TSN-Optical Flow system	94.37
Spatiotemporal method	92.49
TSN-RGB algorithm	90.45

Table 7. AUC_{score} Outcome of EPWOD-POAADP method with existing models under UCSDPed2 dataset.

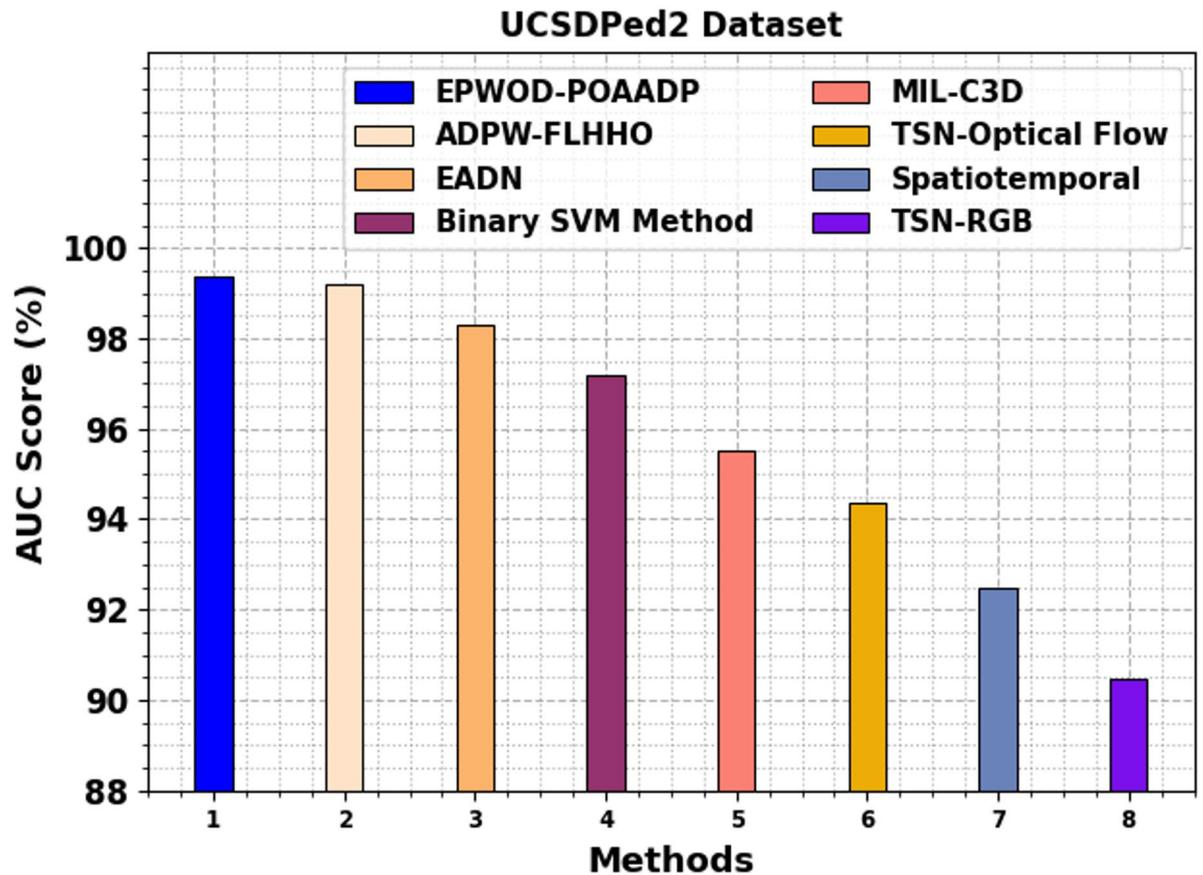


Fig. 12. AUC_{score} outcome of EPWOD-POAADP method under UCSDPed2 dataset.

Methods	CT (sec)
EPWOD-POAADP	8.12
ADPW-FLHHO model	11.34
EADN method	11.23
Binary SVM method	11.76
MIL-C3D technique	12.47
TSN-optical flow system	13.72
Spatiotemporal method	19.52
TSN-RGB algorithm	11.87

Table 8. CT analysis of EPWOD-POAADP technique with existing models under UCSDPed2 dataset.

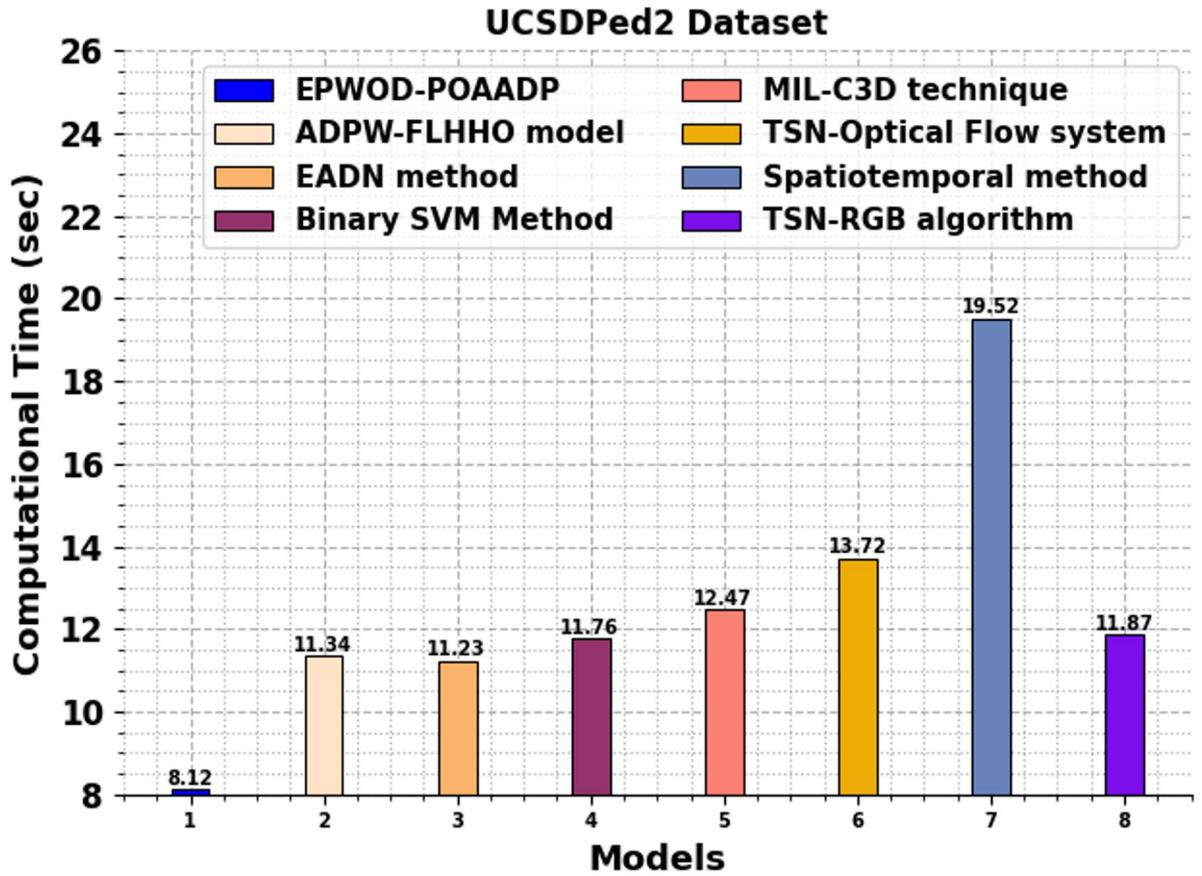


Fig. 13. CT analysis of EPWOD-POAADP technique with existing models under UCSDPed2 dataset.

UCSDPed2 Dataset	
Methods	AUC Score (%)
EPWOD-POAADP	99.35
WNN	98.46
POA	97.71
CapsNet	96.95
Faster R-CNN	96.45
MF	95.81

Table 9. Comparative performance evaluation of the EPWOD-POAADP methodology through ablation under the UCSDPed2 dataset.

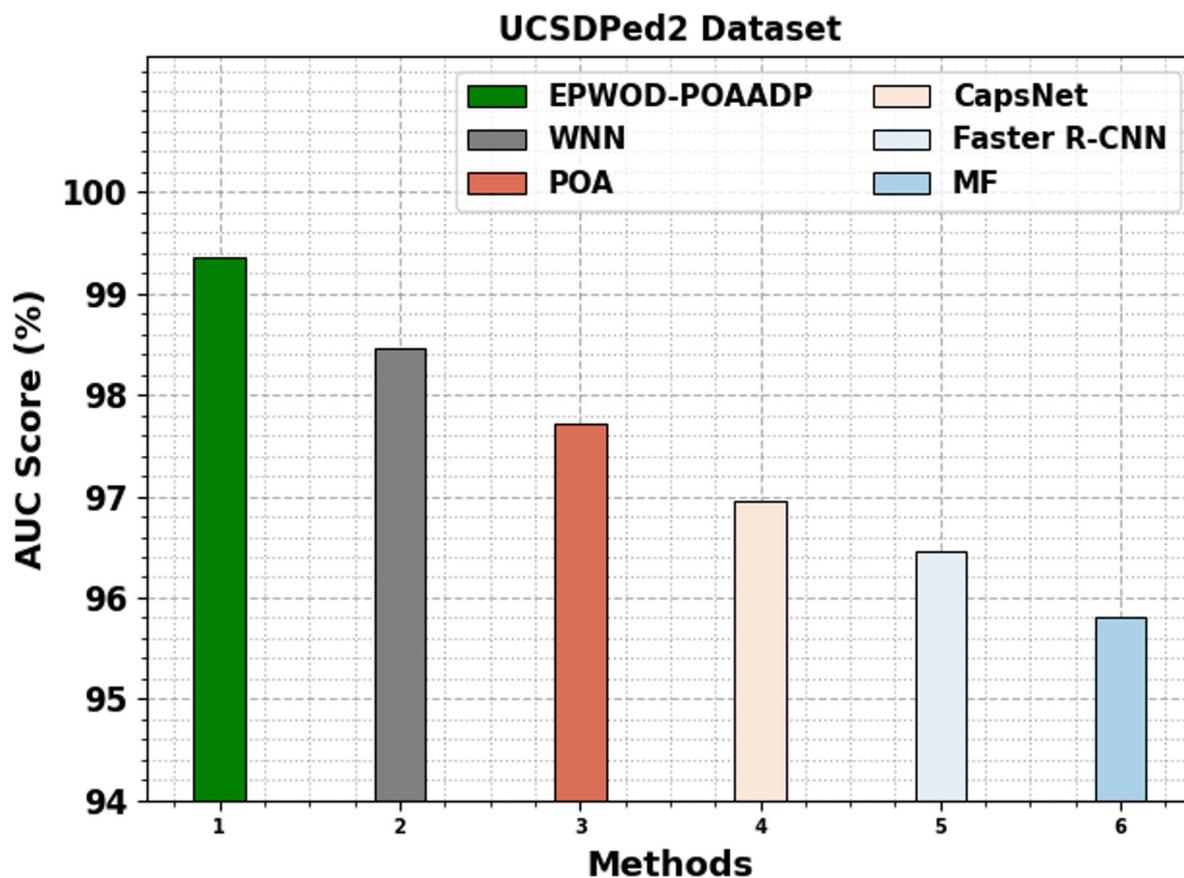


Fig. 14. Comparative performance evaluation of the EPWOD-POAADP methodology through ablation under the UCSDPed2 dataset.

Methods	FLOPs	GPU
Nearest	136.01	2161
Bilinear	135.86	3049
Deconv	143.93	2748
Pixel Shuffle	167.31	2895
Dysample	136.07	3530
CARAFE	135.98	2314
EPWOD-POAADP	90.34	1200

Table 10. Comparison of upsampling methods based on flops and GPU usage.

Data availability

The data supporting this study's findings are openly available at <http://www.svcl.ucsd.edu/projects/anomaly/dataset.html>, reference number [23].

Received: 15 February 2025; Accepted: 8 December 2025

Published online: 10 December 2025

References

1. Campisi, T., Ignaccolo, M., Inturri, G., Tesoriere, G. & Torrisi, V. Evaluation of walkability and mobility requirements of visually impaired people in urban spaces. *Res. Transport. Bus. Manag.* **40**, 100592 (2021).
2. Bentzen, B. L. et al. Wayfinding problems for blind pedestrians at noncorner crosswalks: novel solution. *Transp. Res. Rec.* **2661** (1), 120–125 (2017).
3. Chanana, P., Paul, R., Balakrishnan, M. & Rao, P. V. M. Assistive technology solutions for aiding travel of pedestrians with visual impairment. *J. Rehabilitation Assist. Technol. Eng.* **4**, 2055668317725993 (2017).
4. Frazila, R. B., Zukhruf, F., Simorangkir, C. O. & Burhani, J. T. Constructing pedestrian level of service based on the perspective of visual impairment person. In *MATEC Web of Conferences*. Vol. 270. 03009. (EDP Sciences, 2019).
5. Mattsson, P. et al. Improved usability of pedestrian environments after dark for people with vision impairment: An intervention study. *Sustainability* **12**(3), 1096 (2020).
6. Cohen, A. & Dalyot, S. Route planning for blind pedestrians using openstreetmap. *Environ. Plann. B: Urban Analytics City Sci.* **48** (6), 1511–1526 (2021).
7. Hsieh, I. H., Cheng, H. C., Ke, H. H., Chen, H. C. & Wang, W. J. A CNN-based wearable assistive system for visually impaired people walking outdoors. *Appl. Sci.* **11**(21), 10026 (2021).
8. El-Taher, F. E. Z., Taha, A., Courtney, J. & McKeever, S. A systematic review of urban navigation systems for visually impaired people. *Sensors* **21**(9), 3103 (2021).
9. Mediastika, C. E., Sudarsono, A. S. & Kristanto, L. The sound perceptions of urban pavements by sighted and visually impaired people—A case study in Surabaya, Indonesia. *J. Urbanism: Int. Res. Placemaking Urban Sustain.* **15** (1), 106–129 (2022).
10. Sreeraman, Y. et al. Enhancing anomaly detection in pedestrian walkways using improved sparrow search algorithm with parallel features fusion model. *Fusion Pract. Appl.* **14**(2). (2024).
11. Bhatlawande, S., Dhande, S., Gupta, D., Madake, J. & Shilaskar, S. Pedestrian and vehicle detection for visually impaired people. In *International Conference on Communications and Cyber Physical Engineering 2018*. 37–51. (Springer Nature Singapore, 2023).
12. Kumar, A., Chakravarty, A., Choudhary, A. & Indu, S. Camera-based mobility framework for visually impaired pedestrians in unstructured environments. In *2024 IEEE Intelligent Vehicles Symposium (IV)*. 311–316 (IEEE, 2024).
13. Adi, H. P., Heikoop, R. & Wahyudi, S. I. Enhancing inclusivity: designing disability friendly pedestrian pathways. *Int. J. Saf. Secur. Eng.* **14**(3). (2024).
14. Sharma, S. & George, B. A shoe with bipolar electrodes for ground impedance based pedestrian pathway classification. *IEEE Sens. J.* (2024).
15. Hamadi, A. & Latoui, A. An accurate smartphone-based indoor pedestrian localization system using ORB-SLAM camera and PDR inertial sensors fusion approach. *Measurement* **240**, 115642 (2025).
16. Yoshikawa, T. & Premachandra, C. Pedestrian crossing sensing based on Hough space analysis to support visually impaired pedestrians. *Sensors* **23**(13), 5928 (2023).
17. Guo, X. & Shen, Z. Smart pedestrian crossing design by using smart devices to improve pedestrian safety. *Rev. Adhes. Adhes.* **11**(3). (2023).
18. Ahmed, S. & Islam, S. Methods in detection of median filtering in digital images: A survey. *Multimedia Tools Appl.* **82** (28), 43945–43965 (2023).
19. Zia, H. et al. Advancing road safety: A comprehensive evaluation of object detection models for commercial driver monitoring systems. *Future Transport.*, **5**(1), 2 (2025).
20. Katkam, S., Tulasi, V. P., Dhanalaxmi, B. & Harikiran, J. *Multi-class Diagnosis of Neurodegenerative Diseases using Effective Deep Learning Models with Modified DenseNet-169 and Enhanced DeepLabV3+*. (IEEE Access, 2025).
21. Hu, X. et al. Research on RTD fluxgate induction signal denoising method based on particle swarm optimization wavelet neural network. *Sensors*, **25**(2), 482 (2025).
22. Ajenikoko, G. A., Adebayo, I. G. & Adeleke, B. S. *Hybridization of Mayfly-Pelican Optimization Algorithm for Selection of CNN Optimal Hyper-Parameters*.
<http://www.svcl.ucsd.edu/projects/anomaly/dataset.html>.
24. Alohal, M. A. et al. Anomaly detection in pedestrian walkways for intelligent transportation system using federated learning and Harris hawks optimizer on remote sensing images. *Remote Sens.* **15**(12), 3092 (2023).
25. Li, Z. et al. Self-supervised feature contrastive learning for small weak object detection in remote sensing. *Remote Sens.* **17**(8), 1438 (2025).

Acknowledgements

The authors thank the King Salman Center For Disability Research for funding this work through Research Group no KSRG-2024- 143.

Author contributions

All authors wrote the main manuscript text, all authors prepared all figures, all authors analysis results and all authors reviewed the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to R.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025