International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

Automated Object Detection and Classification using Krill Herd Algorithm with Deep Learning on Surveillance Videos

V. Saikrishnan[1], Dr. M. Karthikeyan[2]

[1]Department of Computer and Information science, Faculty of Science, Annamalai University, Tamil Nadu, India
Corresponding Author Email: *saikrishnan1074[at]gmail.com,*

[2]Department of Computer and Information science, Faculty of Science, Annamalai University, Tamil Nadu, India
Email: *karthiaucse[at]gmail.com*

Abstract: ***In the domain of video surveillances, the implementation of deep learning (DL) for object detection and classification is developed as a game-changer. This paper introduces a comprehensive solution integrating the power of DL methods to handle these fundamental tasks. Leveraging state-of-the-art neural networks (NN), our technique allows consistent object identification and categorization within surveillance videos, providing improved security, real-time context awareness, and enriched decision-making abilities. This study develops an Automated Object Detection and Classification using Krill Herd Algorithm with Deep Learning (AODC-KHADL) technique on Surveillance Videos. The introduced AODC-KHADL method efficiently detects and classifies the objects into numerous categories. This technique starts with the incorporation of YOLO-v5, a recent object detection method popular for its excellent accuracy and speed for identifying objects in videos and images. For enhancing YOLO-v5's detection potential, we utilize Random Vector Functional Link (RVFL) classification, a multipurpose and robust machine learning (ML) approach. In this context, we present the Krill Herd Algorithm (KHA), a nature-inspired optimization method inspired by the collective behavior of krill swarms. By using extensive examination and assessment, we exhibit the model's capability in real-time video surveillance applications. The simulation values of the AODC-KHADL technique are tested on benchmark video and it is emphasized the higher performance of the AODC-KHADL system with other models.***

Keywords: Object classification; Video surveillance; Object detection; Krill Herd Algorithm; Deep learning

1. Introduction

Video surveillance denotes a method of assessing video arrangements and it remnants to be an energetic part of computer vision (CV). It awards huge size of information storage and presentation [1]. There were 3 types of Video surveillance pursuits. Video surveillance actions are completely independent, manual, partially independent, and manual [2]. Physical video surveillance is described as the term that videos could be examined by the humanoid [3]. These classifications were generally utilized where partial independent video surveillance has definite form of videos clarification but with social intervention. Object recognition in video sequence can be a nonstop evolving region which has an excessive amount of submissions in several areas such as remote sensing, biomedical imaging, biometry,robot navigation, vehicle navigation, visual inspection, and video surveillance [4]. Due to the hasty growth and user-friendliness of first- class cameras in video taking skills, video is a low-cost foundation of information [5]. These clues to excessive attention in the object identification and exploration of video categorizations [6]. The various stages present in object identification were object detection, pre-processing, feature extraction, identification depending on features removed and translation of videos in frames. Object identifications in videos denote a complex method that requires extremely precise and healthy methods [7].

Deep learning (DL) depending on object identification approaches can be a kind of CV method that influences deep neural network (DNN) for identifying and then discovering objects in images or videos [8]. These approaches were accomplished in huge volumes of information against categorized images for studying the configurations and features that define several things. The related DL object recognition approaches were reliant on convolutional neural networks (CNN), which have been developed for handling image databases [9]. These methods usually utilize a combination of identification, object application generation and feature extraction to classify objects in the images [10].

The author at [11] proposed for training field approve scene explicit pedestrian sensors in an unofficial method. The general detector was repositioned for several directed arenas in a labeled basis field dataset with the absence of human-interpreted directed cases. In depth, it can primarily extend common sensors to dual boundary identification and composed inflexible specimens as un-labelled directed specimens related to the identification self-reliance. The author [12] proposed a MODT method. Most of the video files are converted into morphological operations as per the extent of frames via region growing methodology. The authors designed a novel technique for movement tracking and object identification. In [13], the author provided an effective video object identification as well as tracking method. The undefined morphological filter were implemented for destroying the sounds, which is accessible in forefront segmented frames.

Jha et al. [14] proposed an N-YOLO approach which slightly than resizing image phase into YOLO, this one separates into permanent sized images manipulated in YOLO. Then it combined detection results with inference outcomes of sub-images at dissimilar intervals by using relationship-oriented tracing process for object identification as well as tracking to reduce the sum of calculation. The author [15], a Background Modelling method has been projected by a Biased Illumination Field Fuzzy C-means technique to classify movable objects exactly. Then, the non- stationary pixel disconnected from stationary pixel by employing Background Subtraction. At last, the Biased Illumination Field Fuzzy- C means method was expanded to enhance divide exactness through gathering within shifting and then sound radiance circumstances.

This study designed an Automated Object Detection and Classification through Krill Herd Algorithm with Deep Learning (AODC-KHADL) technique on Surveillance Videos. The presented AODC-KHADL method efficiently identifies and classifies the objects into numerouscategories. To complement YOLO-v5's detection capabilities, we employ Random Vector Functional Link (RVFL) classification, a powerful and versatile machine learning technique. In this regard, we introduce the Krill Herd Algorithm (KHA), a nature-inspired optimization algorithm inspired by the collective behavior of krill swarms. The simulation values of the AODC-KHADL system are tested on benchmark video and it exhibited the higher performance of the AODC-KHADL system with other approaches.

2. The proposed method

We have proposed an innovative AODC-KHADL system for detecting and classifying objects on surveillance videos. The introduced AODC-KHADL model effectively identifies and classifies the objects into various categories. To accomplish the AODC- KHADL model executes RVFL classification, YOLO-v5 object detection, and KHA-based parameter optimization.

2.1 Object Detection using YOLO-v5

The AODC-KHADL method primarily accomplishes object detection by employing YOLO-v5 framework. Object detection has a major functionin CV that can be an important development with the arrival of YOLO-v5 [16]. YOLO-v5 constitutes the resultant of works to make a remarkably accurate and effective object detection technique. This was attained by employing a restructured model, which directly forecasts class probabilities and bounding boxes, computational complexity and minimizing redundancy. YOLO-v5 has determined extensive applications, from medical imaging and autonomous vehicles to video surveillance, revolutionizing the method we observing and interacting with visual information. Additionally, YOLO-v5 can simply be adaptable, permitting users to fine-tune the process for particular tasks, creating a better selection for those find stability among accuracy and speed in object detection. With their real-time abilities and significant identification effectiveness, YOLO-v5 endures to drive the boundaries of what is probable in the domain of object detection that provides an effective solution for anextensive array of CVdifficulties.

2.2 Object Classification using RVFL Model

The AODC-KHADL algorithm exploited the RVFL system for the purpose of object classification. RVFL networks describe a new and effective method for ML and regression function. RVFL's different feature lies in its unique model that comprises a hidden layer of random feature mappings and a following output layer [17]. In RVFL, the hidden layer implements a group of activation functions, random weights, and biases to provide input data into a high-dimensional feature space. This conversion successfully decorrelated the input data that is specifically beneficial for linearly connected issues. Further, the model employs a linear output layer to execute regression function or binary classification, creating it computationally effectual and particularly simple for training.

The main mathematical expressions guiding the RVFL network can begiven as:

The hidden layer of an RVFL network maps the input data represented as X, into a high-dimensional space applying activation functions (ϕ), random weights (W), andrandom biases (b):

$$H = \phi(X \cdot W + b) \quad (1)$$

Now, H indicates the modified features, and ϕ signifies the activation function (for example, ReLU orsigmoid).

Then the feature conversion, RVFL accomplishes linear regression utilizing the output weights (O)andmodified features H:

$$Y = H \cdot O \quad (2)$$

In case Y is the predicted output.

To avoid overfitting, RVFL generally contains a regularization term, normally L2 regularization that could be includedin the loss function:

$$Loss = \frac{1}{2N}\sum_{i=1}^{N}(Y_i - H_i \cdot O)^2 + \frac{\lambda}{2}\|O\|_2^2 \quad (3)$$

In this equation, N denotes the number of data points in the training set and λ controls the strength of regularization.

RVFL networks provide numerous benefits including their capability for effectively controlling higher- dimensional data, their difficulties with overfitting, and simple of deployment. These features create RVFL a useful tool in diverse regression and ML applications. Fig. 1 depicts the architecture of RVFL.

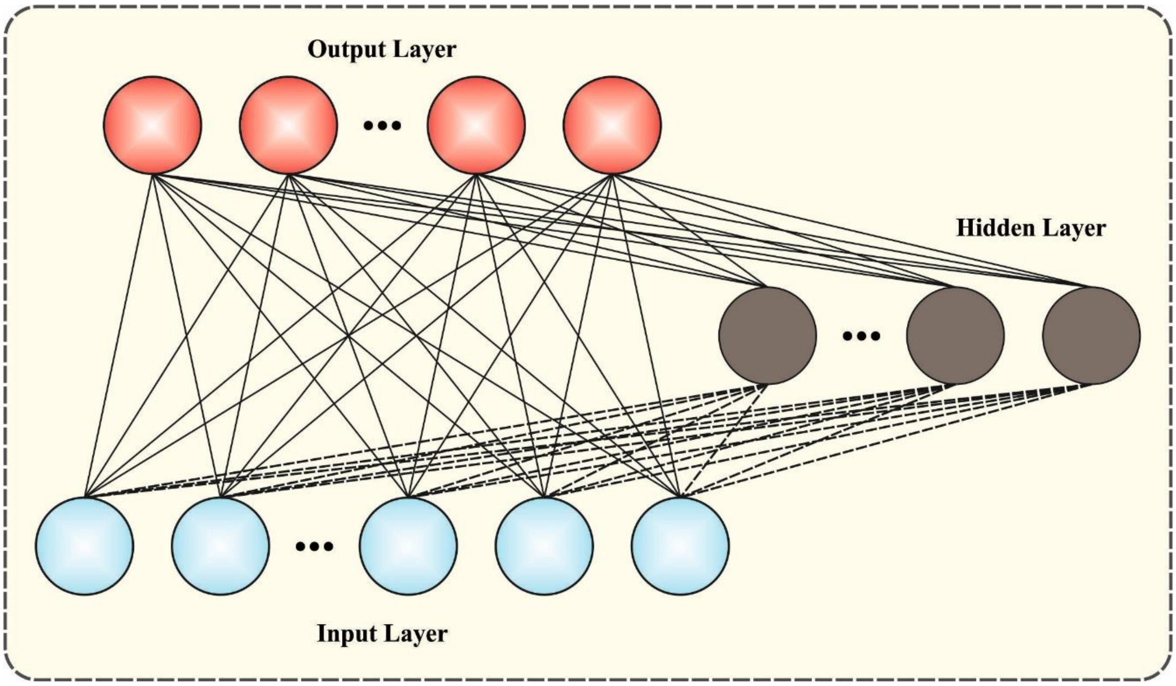

Figure 1: Structure of RVFL

2.3 Parameter Tuning using KHA Algorithm

The KHA method has been employed as a parameter optimization technique for increasing efficiency of detection.KHA is a metaheuristic approach that simulates the for aging of krill population and density-dependent attraction [18]. Three operations define the krill movement in search range. The initial one is the movement affected by the other individual krill's and is computed by Eq. (4):

$$N_i^{t+1} = N^{\max} \alpha_i + \omega_i N_i^t \quad (4)$$

Where, $\alpha_i = \alpha_i^{local} + \alpha_i^{target} . \alpha_i^{local}$ denotes the local effect generated by the neighbors and α_i^{target} shows the direction defined by the optimal krill individuals. $N^{\max}$ indicates the maximal induced speed, ω_i indicates the inertia weight of induced movement within [0,1], and N_i^t shows the last induced movement. The foraging motion is the second operation that can be evaluated according to the food position and prior knowledge as follows:

$$F_i^{t+1} = V_f \beta_i + \omega_f F_i^t \quad (5)$$

Now $\beta_i = \beta_i^{food} + \beta_i^{best}$ which β_i^{food} denotes the attractiveness of food and β_i^{best} indicates the impact of better fitness. V_f shows the speed of foraging and ω_f indicates the inertia weight for foraging movement within [0,1]. The krill individual moves randomly in the searching region at the last operation that is named physical diffusion:

$$D_i^{t+1} = D^{\max} \delta \quad (6)$$

Here δ is a random vector, and $D^{\max}$ denotes the maximum diffusion speed. Using the abovementioned motion functions, the final movement of krill can be computed by using Eq. (7):

$$x_i^{t+\Delta t} = x_i^t + \Delta t(N_i^{t+1} + F_i^{t+1} + D_i^{t+1})$$

$$\Delta t = C_t \sum_{j=1}^{n} (U_j - L_j) \quad (7)$$

Now n denotes the dimension of searching range, U_j and L_j indicates the upper and lower boundaries of the j^{th} paramter. C_t shows the constant within [0,2]. Lastly, to improve the performance, genetic operators including crossover and the mutation are performed. Each step forms a KHA iteration which continue until ending condition is satisfied.

3. Result and Discussion

The performance analysis of the AODC-KHADL model is tested on the UCSDPed2 datasets [19] that comprise 2 subsets namely pedestrian 1 and pedestrian 2 datasets. Table 1 represents the database description. Fig. 2 exhibits the sample test image with corresponding ground truth images.

Table 1 Description of dataset

Dataset	Testbed	Frames No.	Time (sec)
UCSDped2	Pedestrian1 Dataset	360	12
	Pedestrian2 Dataset		

Figure 2: a) Original Image b) Ground Truth Image

Table 2 and Fig. 3 considers the average detection accuracy of the AODC-KHADL method on 2 databases. The figure shows that the AODC-KHADL technique obtained better performance over the other techniques on 2 sub-datasets. Based on surveillance ped1 dataset, the AODC-KHADL system provided an improved average accuracy of 98.21% while the CIHSART-ODT, DLA-DT, Region-CNN, and FR-CNN methods acquired lesser average accuracy of 98%, 97%, 97%, and 85%. Besides, with surveillance ped2 database, the AODC-KHADL model has offered higher average accuracy of 92.92% whereas the CIHSART-ODT, DLA-DT, Region-CNN, and FR-CNN techniques obtained a minimum average accuracy of 91%, 90%, 87%, and 82% respectively.

Table 2: Average $accu_y$ outcome of Analysis of AODC-KHADLmethod on two datasets

Methods	AODC-KHADL	CIHSART-ODT	DLA-DT	Region-CNN	FR-CNN
Surveillance Ped-1	98.39	98.00	97.00	97.00	85.00
Surveillance Ped-2	93.52	91.00	90.00	87.00	82.00

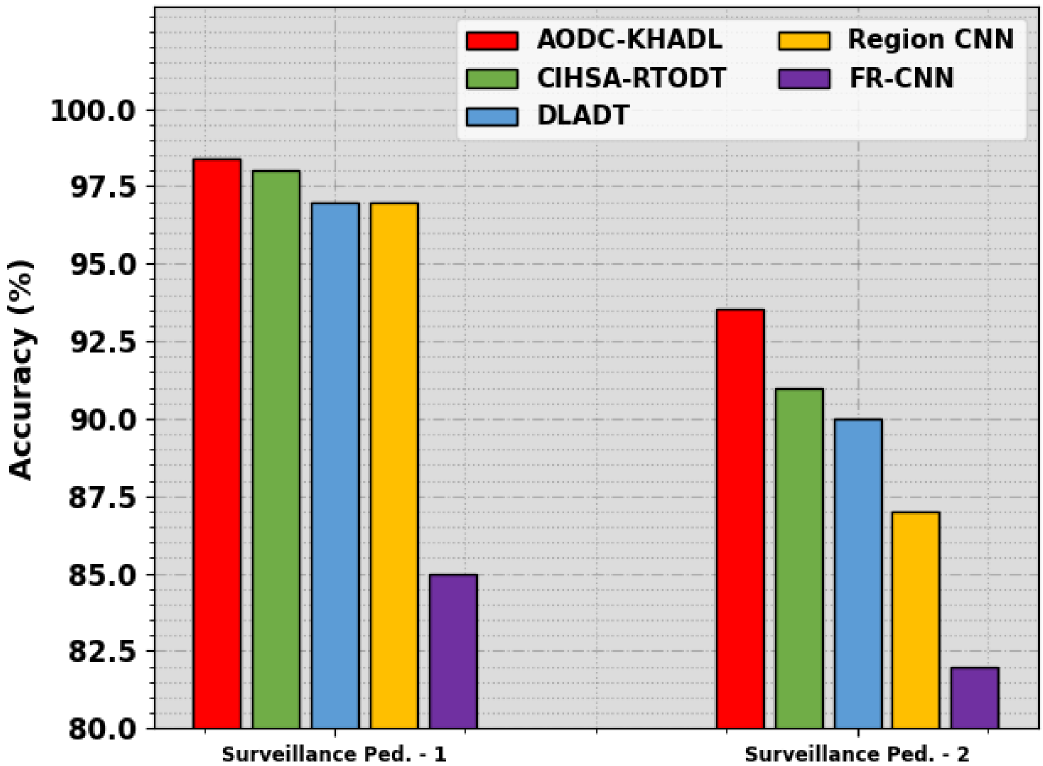

Figure 3: Average accuracy of AODC-KHADL model on two sub-datasets

Table 3 and Fig. 4 illustrate the AUC of the AODC-KHADL technique on 2 databases [6, 20-22]. The figure denotes that the AODC-KHADL method achieved improved performance over the other models on 2 sub-datasets. Additionally, with surveillance ped1 dataset, the AODC-KHADL method has offered greater AUC of 98.18%

whereas the MP-PCA, SF, SFMP-PCA, M-DT, A-MDN, AD-VAE, and CIHSART-ODT methodologies acquired a decreased value AUC of 61.01%, 66.74%, 67.25%, 82.05%, 91.71%, 95.39%, and 97.12%. Also, with surveillance ped2 database, the AODC-KHADL approach gives a maximum AUC of 95.25% but, the MP-PCA, SF, SFMP-PCA, M-DT, A-MDN, AD-VAE, and CIHSART-ODT systems acquired lessening value AUC of 69.92%, 55.96%, 61.33%, 82.99%, 91.25%, 92.47%, and 93.92% corresponding.

Table 3: AUC analysis of AODC-KHADL model on two sub datasets

Models	Surveillance Ped1	Surveillance Ped2
MP-PCA	61.01	69.92
SF	66.74	55.96
SFMP-PCA	67.25	61.33
M-DT	82.05	82.99
A-MDN	91.71	91.25
AD-VAE	95.39	92.47
CIHSART-ODT	97.12	93.92
AODC-KHADL	98.18	95.25

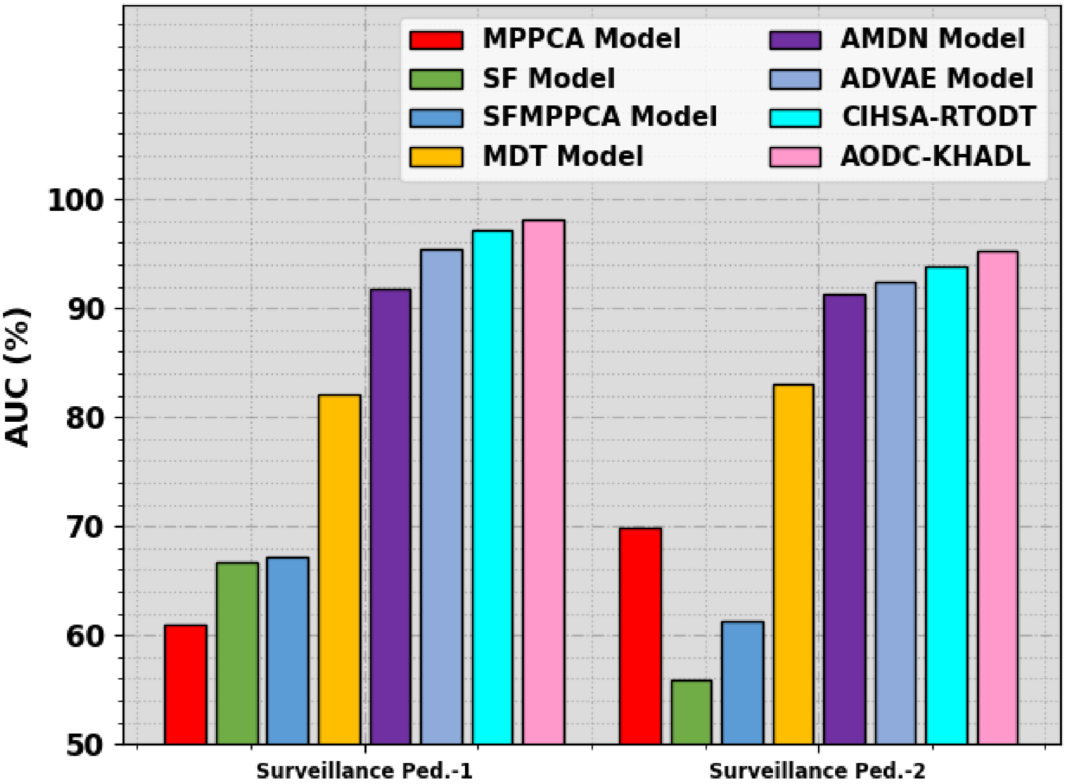

Figure 4: AUC analysis of AODC-KHADL model on two sub datasets

Table 4 and Fig. 5 exhibit a running time (RT) performance of the AODC-KHADL technique over the other models. The simulated values show that the AODC-KHADL method has offereda diminished RT over the other models. According to surveillance ped-1 dataset, the AODC-KHADLsystemgivesa lower RT of accuracy of 2.01s while the M-DT, SCLF, A-MDN, AD-VAE, and CIHSART-ODT methodologies obtained a decreased value RT of 20.61s, 20.11s, 11.73s, 3.94s, and 2.67s individually. Meanwhile, on surveillance ped-2 dataset, the AODC-KHADL technique gives least RT of accuracy of 2.40s whereas the M-DT, SCLF, A-MDN, AD-VAE, and CIHSART-ODT algorithms acquire lessen RT of 22.94s, 18.48s, 13.02s, 6.16s, and 3.98s appropriately.

Table 4: RT analysis of AODC-KHADL model with other approaches on two sub-datasets

Methods	Pedestrian1	Pedestrian2
M-DT	20.61	22.94
SCLF	20.11	18.48
A-MDN	11.73	13.02
AD-VAE	3.94	6.16
CIHSART-ODT	2.67	3.98
AODC-KHADL	2.01	2.40

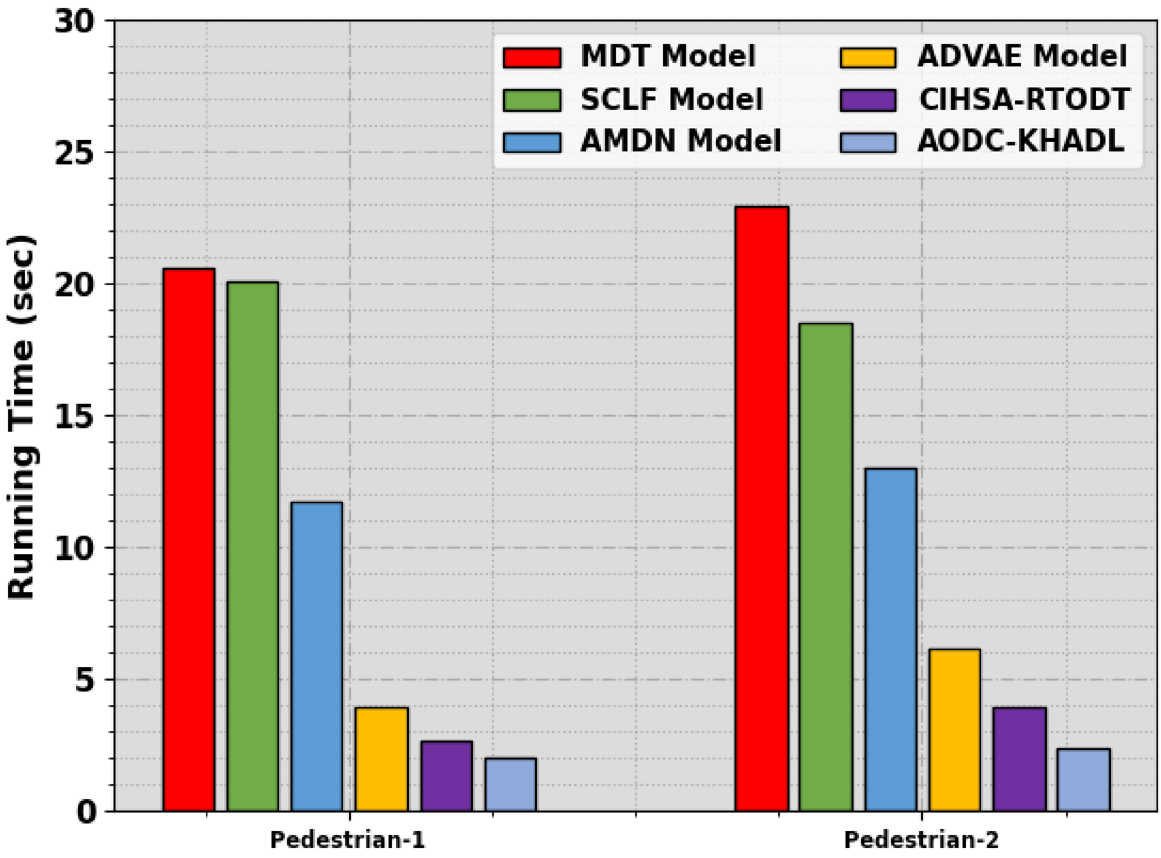

Figure 5: RT analysis of AODC-KHADL model on two sub datasets

4. Conclusion

In this work, we have developed an innovative AODC-KHADL technique for detecting and classifying objects on surveillance videos. This introduced AODC-KHADL system effectively identifies and classifies the objects into various categories.To complement YOLO-v5's detection capabilities, we employ RVFL classification, a powerful and versatile ML technique. In this regard, we introduce the KHA, a nature-inspired optimization algorithm inspired by the collective behavior of krill swarms. The simulated values of the AODC-KHADL approach are tested on benchmark video and it is exhibited the improved performance of the AODC-KHADL model with other approaches.

References

[1] Chandrakar, R., Raja, R., Miri, R., Sinha, U., Kushwaha, A.K.S. and Raja, H., 2022. Enhanced the moving object detection and object tracking for traffic surveillance using RBF-FDLNN and CBF algorithm. *Expert Systems with Applications*, *191*, p.116306.

[2] Luo, X., Wang, Y., Cai, B. and Li, Z., 2021. Moving object detection in traffic surveillance video: new MOD-AT method based on adaptive threshold. *ISPRS International Journal of Geo-Information*, *10*(11), p.742.

[3] Tom, A.J. and George, S.N., 2020. Simultaneous reconstruction and moving object detection from compressive sampled surveillance videos. *IEEE Transactions on Image Processing*, *29*, pp.7590-7602.

[4] Ingle, P.Y. and Kim, Y.G., 2022. Real-Time Abnormal Object Detection for Video Surveillance in Smart Cities. *Sensors*, *22*(10), p.3862.

[5] Kim, J.H., Choi, J.H., Park, Y.H. and Nasridinov, A., 2021. Abnormal situation detection on surveillance video using object detection and action recognition. *Journal of Korea Multimedia Society*, *24*(2), pp.186-198.

[6] Alotaibi, M.F., Omri, M., Abdel-Khalek, S., Khalil, E. and Mansour, R.F., 2022. Computational intelligence-based harmony search algorithm for real-time object detection and tracking in video surveillance systems. *Mathematics*, *10*(5), p.733.

[7] Joy, F. and Vijayakumar, V., 2022. An improved Gaussian Mixture Model with post-processing for multiple object detection in surveillance video analytics. *International journal of electrical and computer engineering systems*, *13*(8), pp.653-660.

[8] Arunnehru, J., 2021. Deep learning-based real-world object detection and improved anomaly detection for surveillance videos. *Materials Today: Proceedings*.

[9] Elhoseny, M., 2020. Multi-object detection and tracking (MODT) machine learning model for real-time video surveillance systems. *Circuits, Systems, and Signal Processing*, *39*, pp.611-630.

[10] Kumar, C. and Punitha, R., 2020, August. Yolov3 and yolov4: Multiple object detection for surveillance applications. In *2020 Third international conference on smart systems and inventive technology (ICSSIT)* (pp. 1316-1321). IEEE.

[11] Mou, Q., Wei, L., Wang, C., Luo, D., He, S., Zhang, J., Xu, H., Luo, C. and Gao, C., 2021. Unsupervised

domain-adaptive scene-specific pedestrian detection for static video surveillance. *Pattern Recognition*, *118*, p.108038.

[12] Elhoseny, M., 2020. Multi-object detection and tracking (MODT) machine learning model for real-time video surveillance systems. *Circuits, Systems, and Signal Processing*, *39*, pp.611-630.

[13] Mahalingam, T. and Subramoniam, M., 2021. A robust single and multiple moving object detection, tracking and classification. *Applied Computing and Informatics*, *17*(1), pp.2-18.

[14] Jha, S., Seo, C., Yang, E. and Joshi, G.P., 2021. Real time object detection and trackingsystem for video surveillance system. *Multimedia Tools and Applications*, *80*, pp.3981-3996.

[15] Kalli, S., Suresh, T., Prasanth, A., Muthumanickam, T. and Mohanram, K., 2021. An effective motion object detection using adaptive background modeling mechanism in video surveillance system. *Journal of Intelligent & Fuzzy Systems*, *41*(1), pp.1777-1789.

[16] Katsamenis, Iason, Eleni Eirini Karolou, Agapi Davradou, Eftychios Protopapadakis, Anastasios Doulamis, Nikolaos Doulamis, and Dimitris Kalogeras. "TraCon: A novel dataset for real-time traffic cones detection using deep learning." *In Novel & Intelligent Digital Systems: Proceedings of the 2nd International Conference (NiDS 2022)*, pp. 382-391. Cham: Springer International Publishing, 2022.

[17] Tang, L., Wu, Y. and Yu, L., 2018. A non-iterative decomposition-ensemble learning paradigm using RVFL network for crude oil price forecasting. *Applied Soft Computing*, *70*, pp.1097-1108.

[18] Kaya, E., Baştemur Kaya, C., Bendeş, E., Atasever, S., Öztürk, B. and Yazlık, B., 2023. Training of Feed-Forward Neural Networks by Using Optimization Algorithms Based on Swarm-Intelligent for Maximum Power Point Tracking. *Biomimetics*, *8*(5), p.402.

[19] http://www.svcl.ucsd.edu/projects/anomaly/dataset.html

[20] 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. *Safety Science*, *142*, p.105356.

[21] Xu, M., Yu, X., Chen, D., Wu, C. and Jiang, Y., 2019. An efficient anomaly detection system for crowded scenes using variational autoencoders. *Applied Sciences*, *9*(16), p.3337.

[22] Murugan, B.S., Elhoseny, M., Shankar, K. and Uthayakumar, J., 2019. Region-based scalable smart system for anomaly detection in pedestrian walkways. *Computers & Electrical Engineering*, *75*, pp.146-160.