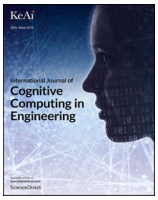
International Journal of Cognitive Computing in Engineering 5 (2024) 343–356

Contents lists available at ScienceDirect

International Journal of Cognitive Computing in Engineering

journal homepage: www.keaipublishing.com/en/journals/international-journal-of-cognitive-computing-in-engineering/



An effective object detection and tracking using automated image annotation with inception based faster R-CNN model

K. Vijiyakumar, Assistant Professor [a,*], V. Govindasamy, Associate Professor [b], V. Akila, Assistant Professor [c]

[a] Department of Computing Technologies, School of computing, SRM institute of Science and Technology, Kattankulathur campus, Chengalpattu district, Tamilnadu, India
[b] Department of Information Technology, Pondicherry Engineering College Pondicherry, India
[c] Department of Computer Science and Engineering, Pondicherry Engineering College, Pondicherry, India

ARTICLE INFO

Keywords:
Object detection
Tracking
Convolutional neural network
Inception v2
Image Annotation

ABSTRACT

The present study advances object detection and tracking techniques by proposing a novel model combining Automated Image Annotation with Inception v2-based Faster RCNN (AIA-IFRCNN). The research methodology utilizes the DCF-CSRT model for image annotation, Faster RCNN for object detection, and the inception v2 model for feature extraction, followed by a softmax layer for image classification. The proposed AIA-IFRCNN model is evaluated on three benchmark datasets: Bird (Dataset 1), UCSDped2 (Dataset 2), and Under Water (Dataset 3), to determine prediction accuracy, annotation time, Center Location Error (CLE), and Overlap Rate (OR). The experimental results indicate that the AIA-IFRCNN model outperformed existing models regarding detection accuracy and tracking performance. Notably, it achieved a maximum detection accuracy of 95.62 % on Dataset 1, outperforming other models. Additionally, it achieved minimum average CLE values of 4.16, 5.78, and 3.54, and higher overlap rates of 0.92, 0.90, and 0.94 on the respective datasets (1, 2 and 3). Hence, this research work on object detection and tracking using the AIA-IFRCNN model is essential for improving system efficiency and fostering innovation in the field of computer vision and object tracking.

1. Introduction

Visual Object Tracking (VOT) is identifying a random destination, represented by a Region of Interest (ROI), in a video. Different application areas namely security and surveillance, multimedia, robotics, augmented reality, and even entertainment (Deori & Thounaojam, Jul., 2014), depends upon the tracking process. Despite recent advancements in the field, Convolutional Neural Networks (CNNs), tracking remains an essential operation in the domain of computer vision. In general, tracking algorithms construct a process of the destination ROI and use the resulting process to seek targets in subsequent frames. The challenges originate initially from the discriminative capabilities of the target model, which must be unique such that the target does not drift to identical objects, while the rest of the target is adaptive enough to the target's inter-frame look changes. A tracking algorithm should also handle occlusions, fast target motions, and out-of-view scenes. Another important aspect is the tracker speed, which is the time it takes to locate the target in each frame. Real-world requirements constrain the trackers per-frame processing time.

Recently, CNNs have been more effective in alternate Computer Vision tasks, like image classification, object prediction as well as semantic segmentation (Schmidhuber, 2015). The semantically relevant representation extracted from visible details is responsible for the effectiveness. As a result, CNNs are used to monitor operations. In general, CNN-based trackers filter convolutional feature-based targets from consecutive frames and then cross-correlating the target method by frame technique to locate the target. It is based on online network parameter optimization or offline learning to develop discriminative features while monitoring faster tracking speed on Graphics Processing Units (GPUs). To report the challenges required by smart functions, like embedded devices and robots (Kamate & Yilmazer, 2015), a tracking model should force the speed-accuracy trade-off to a greater extent.

For CNN-relied trackers, it represents the present neural structures at the time of keeping the count of layers comparatively minimum. The maximum number of annotated video datasets is used like ImageNet VID or TrackingNet (Müller et al., 2018) are used for training. Before

* Corresponding author.
E-mail address: vijiya.kumar@gmail.com (K. Vijiyakumar).

https://doi.org/10.1016/j.ijcce.2024.07.006
Received 25 July 2022; Received in revised form 24 July 2024; Accepted 29 July 2024
Available online 3 August 2024

applying the traditional features, histogram-aided descriptors are applied like Histogram of Oriented Gradients (HOG), descriptors in the Discriminative Correlation Filter (DCF) methodologies (Henriques et al., 2015) or color histograms in Mean Shift (MS) model (Comaniciu & Ramesh, 2000). Based on the encoded features, histograms are highly effective for making modifications whereas it is robust computationally. According to the proficiency of histogram-based trackers, and effective representations are obtained by Deep CNNs (DCNN), and two models are combined as a single neural structure, in which histograms are obtained from deep convolutional features, for applying optimal variables.

The details of object position find helpful when required to deduce high level details and are utilized to minimize the operations. The identification of object location from an input image could be attained by two processes (Comaniciu et al., 2003). The first is object recognition, and the second is object tracking. In the first scenario, feature extraction is performed on the image to determine the type relevant to the class of the object based on previous information. The object may be recognized using the learned approach after providing the image. As a result, machine learning (ML) techniques are commonly used for recognition (Goyal et al., 2022; Wahab et al., 2023). However, instead of searching the object classes, the pixel details in the RoI are examined in the next instance, and the area with the most extra ordinary likeness is searched for the recently provided input image frames. Thus, object detection relates to the process of discovering a formerly identified object in the applied input image, in contrast object tracking relates to the task of searching for an object through morphological relationship among nearby frames in the video (Grabner et al., 2010). Although, the objects that exist in the synthetic or real-time images can be tracked under varying image quality, resolution, backdrop, and so on (Ahn & Shin, 2018). Thus, the above-mentioned traditional tracking methods may not guarantee a higher detection rate in every scenario.

This paper develops an effective DL-based object detection and tracker model utilizing Automated Image Annotation with Inception v2 based Faster RCNN (AIA-IFRCNN) model. The AIA-IFRCNN model comprises a novel image annotation tool utilizing Discriminative Correlation Filter (DCF) with Channel and Spatial Reliability tracker (CSR) named the DCF-CSRT model. The AIA-IFRCNN model makes uses Faster RCNN as an object detector and tracker, including Region Proposal Network (RPN) and Fast R-CNN. In addition, the inception v2 approach is utilized as the feature extractor to generate the feature map. Finally the softmax layer is utilized to classify images. An Extensive set of experiments is carried out for verify the proficient tracking performance of the AIA-IFRCNN model and the results are investigated under different dimensions. The AIA-IFRCNN model combines of Automated Image Annotation with the Inception-based Faster RCNN architecture, allowing for efficient object detection and tracking by leveraging image annotation and deep learning techniques. It automates the image annotation process, reducing manual intervention and annotation time while ensuring consistency. The Inception v2 model enables effective feature extraction from input images, enhancing accurate object detection and tracking. The model underwent rigorous training and optimization to fine-tune parameters for improved performance on benchmark datasets regarding detection accuracy, center location error (CLE), and overlap rate (OR). Its performance was evaluated against three benchmark datasets: Bird, UCSDped2, and Under Water, quantitatively assessing improvements over existing methods. The innovative architecture, process automation, effective feature extraction, rigorous training and optimization, and thorough benchmarking collectively contributed to the superior performance of the AIA-IFRCNN model in object detection and tracking tasks.

The organization of the paper is as follows, Section 2 displays the various processes involved in the proposed tracker model (DCF-CFRT), and Section 3 showed on the performance validation using 3 datasets namely, Bird (Dataset 1), UCSDped2 (Test004) (Dataset 2), Under Water (Blurred & Crowded) (Dataset 3) and detailed result analysis (i.e. prediction accuracy, annotation time, Center Location Error (CLE) and Overlap Rate (OR)) using 3 datasets and its comparison with other existing models, Section 4 summarizes the conclusions of the work.

2. Related works

Several works have been completed in to combine object recognition and tracking models. This is known as tracking by detection. Massive current trackers have employed and the practical implications obtained by CNNs have been achieved with maximum accuracy. When compared with the previous works applying CNNs are Generic Object Tracking Using Regression Networks (GOTURN) (Held et al., 2016) and Fully Convolutional Siamese Tracker (SiamFC) (Bertinetto et al., 2016) that applies shallow networks and is implemented in concurrent applications. GOTURN applies a convolutional network to extract features from the target and explore areas. Features obtained from areas are integrated with the application of FC layers and associated. The network undergoes training offline, in case of regression, where it predicts the place and size of a target in explored the area. Recently, anchor-based ROI selection has been embedded into a Siamese structure for tracking, which was developed by SiamRPN (Li et al., 2018). The key objective of applying anchors is to map the bounding box of a target where the tracker is suitable to manage the aspect ratio modifications, In contrast classical trackers deal with size alterations when retaining a static aspect ratio. In Murugan et al. (Murugan et al., 2019), an efficient Region based Scalable Convolution Neural Network (RS-CNN) model was projected detect anomalies in pedestrian walkways. It efficiently recognizes the anomalies at an earlier rate and carries out well with the scalability problem. Any other methods like a Mixture of Dynamic Texture (MDT) (Chan & Vasconcelos, 2008), Mixture of optical flow (Kim & Grauman, 2009), Circulant Structure Kernel (CSK) (Henriques et al., 2012), Fast Compressive Tracking (FCT) (Zhang et al., 2014), Discriminative Scale Space Tracker (DSST) (Danelljan et al., 2014), Convolutional Features (CF2) (Ma et al., 2019), and Kernelized Correlation Filter (KCF) tracker (Henriques et al., 2015) made to anomaly detection has been proposed in the literature. In Mehran et al. (Mehran et al., 2009), an efficient Social Force (SF) method has been established for detecting and localization of abnormal performance in crowded videos. A frame in the videos is classified into normal and abnormal using a bag of words process. Anitha et al. (Ramachandran & Sangaiah, 2021) thoroughly assessed of the literature on object identification and tracking with Unmanned Aerial Vehicles (UAVs) for various applications. Payal et al. (Mittal et al., 2022) proposed a quadcopter-based Simulink flight control model to generate a simulated dataset. A Few adjustments have been made to the standard model in order to capture drone footage for object detection while operating in a simulated environment, including pedestrians, other drones, and obstacles.

Long Chen et al. (Chen et al., 2020) introduced the Sample-WeIghted hyPEr Network (SWIPENet), which considers the problem of heterogeneous noise. The network generates a high-resolution, semantically rich feature map. While detecting, the network may discover an undesired object due to difficulty with noisy images. Yang et al. (Yang et al., 2021) presented a solution for real-time object detection utilizing YOLOv3 in 2020. The work outperforms Faster-RCNN in terms of speed and mean Average precision. The process of detecting obstructed objects is slow. Chongwei Liu et al. (Liu et al., 2022) devised a method for detecting sea cucumbers, sea urchins, and scallops underwater. The Generative Adversarial Network is employed to eliminate class imbalance, while AquaNet is proposed for the efficient detecting of small objects. In the year 2020, the Single-Shot Detector (SSD) will be able to tackle the challenge related to small object detection. Hu et al. (Hu et al., 2020) presented a cross-level fusion network to boost feature extraction ability,. The work combines the ResNet and SSD concepts to perform better, outperforming SSD by 7.6 %. The network does not consider the problem of heterogeneous noise. Some of the related works were given in Tables 1 and 2.

Table 1
Stages of development of models in object detection and tracking.

Stages of development	Method	Significance	Thought
Traditional manual features	Manual feature Selection technique using, Machine Learning	Emancipate workers and generate fresh ideas for industrial development	Multi-stage processing technique, sliding window detection, features extraction algorithm, multi-scale pyramid
Deep learning period	CNN	Since then, the use of deep learning in identification has grown fast	Training the network on the data set for extracting the image features
	AlexNet	LeNet model was improved to increase the efficiency	Convolution followed by full join
	VGG16	Improvements were made to AlexNet model	Large convolution kernels were replaced by Several convolution kernels of 3×3.
	Faster RCNN	The models like R-CNN and SPPNet are enhanced. Fast R-CNN improved for realization of deep learning detection.	Candidate region generation network
	YOLO	For the applications of the projects in real time, YOLO model was used	The whole picture was the input and propagated forward at once

3. The proposed tracker

Initially, the DCF-CFRT model annotates the objects that exist in the image. Next, Faster RCNN is applied as an object detector, including the inception v2 model as the shared CNN. Finally, the softmax layer based classification process is carried out.

3.1. DCS-CSRT model

At first, the input videos are changed into a sequence of frames, and the objects in the frame are marked as an objects in the creation of record files. The automated DCS-CSRT method is implemented as image annotation devices to annotate the objects occur in the input frame. It allows for annotating the objects in a single frame and generates automated annotation of the objects in each frame that exist in the video.

3.2. Faster RCNN

A faster Region with Convolutional Neural Network (R-CNN) is a kind of CNN that has evolved from R-CNN. According to the area proposal network, a Faster R-CNN selects an arbitrary area of the image as the proposal area, and trains for obtaining the equivalent type and location of specific in the image. So, it takes high feature recognition of the insulator failure and bird nests from the massive aerial images (Lei & Sui, 2019). Compared with the conventional selective explore technique, the Faster R-CNN is breaking the blockage issue of massive cost in the calculation as the RPN creates equivalent proposal sites. So, the practical examination becomes feasible. Additionally, using an adaptive scale pooling layer, a Faster R-CNN can adapt to arbitrary images and adjust the whole network to enhance the accuracy of deep network recognition. A faster R-CNN model is capable of breaking the time blockage of computation, and ensures an effective prediction rate is achieved. Thus, a Faster R-CNN analysis model is presented to process the feature extraction process in the insulator and the nest for identifying the destination. A Faster R-CNN technique is comprised of 2 CNN

Table 2
Comparison of some of the reported Faster RCNN techniques for object detection.

Reported work	Method	Author's findings
Gavrilescu et al. (2018)	*"Faster RCNN: an Approach to Real-Time Object Detection"*	Experimental results show that 98 % mean average precision is achieved using Faster RCNN algorithm.
Sommer et al. (Sommer et al. 2018)	*"Search Area Reduction Fast RCNN for Fast Vehicle Detection in Large Aerial Imagery"*	Faster RCNN has shown 97.4 % average precision to detecting vehicles.
Irisa & France, (2018)	*"Buried Object Detection from B-Scan Ground Penetration RADAR Data Using Faster RCNN"*	Results show that Faster RCNN has achieved 89 % accuracy on simulated data and also reached 100 % accuracy in some cases.
Manana et al. (2018)	*"Pre-processed Faster RCNN for Vehicle Detection"*	Experiment results proved that almost 100 % accuracy was achieved in detecting vehicles by using Faster RCNN algorithm.
Wang et al. (2017)	*Scene Text Recognition Algorithm Based on Faster RCNN*	In this study, the authors concluded that using Faster RCNN, the average text recognition accuracy rate has reached 90.5 %.
Ning et al. (2017)	*Inception Single Shot Multibox Detector for Object Detection*	The results of this study show that almost 87.8 % accuracy is possible in detecting a person, a car, a tree, a stone and a chair.
Chen et al. (2018)	*Vehicles Detection on Expressway Via Deep Learning: Single Shot Multibox Detector*	In this article, the authors have stated that an 85.8 % detection score was achieved when vehicles were detected on expressway.
Sangari et al. (2023)	*RCNN (Region-based Convolutional Neural Network) and correlation filter tracking algorithm (CFTA) method*	Outperforming the current models, the suggested model yielded the best results in this manner, having an accuracy of 97.89 %.
Aruna et al. (2023))	*CNN—Only Look Once VGG19 algorithm*	2500 photos were utilized for verification and 8200 images were used to train the suggested approach. With its improvements, the VGG19 model achieves 98 % accuracy.
Serdà & Burguera, (2023)	*Object Detection Neural Networks (Faster R-CNN, YOLOv5 and Mask R-CNN).*	Researchers have used the SORT algorithm to accomplish fish tracking using the output of YOLOv5s. Their implementation reduced the frame rate in half but produced decent results. Given the great performance of YOLOv5, even with this significant frame rate drop, fish tracking using a live video feed is still possible.
Sun et al., (2023)	*Target detection algorithm is studied based on PP-YOLO*	while PP-YOLO's average mAP is somewhat lower than YOLOv3, it still deploys models very quickly and tracks well in real time.
Ramachandran & Sangaiah, (2021)	*UAV & GPU platform- IoT agriculture*	This research presents a secured onboard object detection system in precision agriculture.
Li & Lima, (2021)	*Deep residual network ResNet-50*	The approach suggested in this research has excellent precision and good identification effectiveness in terms of mean recognition accuracy, according to the

(*continued on next page*)

Table 2 (*continued*)

Reported work	Method	Author's findings
		experimental findings and data set validation.
Liang et al. (2020)	*Small Object Detection in Unmanned Aerial Vehicle Images Using Feature Fusion and Scaling-Based Single Shot Detector with Spatial Context Analysis*	In this study, the authors have proposed feature fusion and scaling based SSD for detecting small objects. It was concluded that their results have showed superior detection accuracy in detecting small objects when compared with other state-of-the arts methods.
Chandan et al. (2018)	*Real Time Object Detection and Tracking Using Deep Learning and OpenCV*	In this paper, the experiment results show that 99 % confidence level was achieved in detecting various objects classes such as dog, train, person, potted plant etc.
Krishnan et al. (2022)	*Under water object recognition and tracking using hybrid techniques (RetinaNet & EfficientNet) (HDCNN-UODT)*	Author used brackish and URPC datasets for underwater object recognition and tracking. HDCNN-UODT model achieved 94.85 % for identification of "CRAB" object using URPC dataset.

networks, as shown in Fig. 1 (faster-r-cnn-for--object-detection-a-technical-summary-474c5b857b46 @, towardsdatascience.com).

From Fig. 1, the input image was processed through 2 CNN networks namely 1.) Fast R-CNN recognition network present in upper half of the flow diagram, 2.) Regional Proposal Networks (RPN) are in the flow diagram lower half. An RPN samples the arbitrary area of an image as the proposal regions and undergoes training to define with a destination. Furthermore, the Fast R-CNN prediction system is again processed using the data collected by the RPN system, which defines the destination type in the area, and precisely alters the size of the region to locate the particular location of the destination in the image. Initially, the parameter of the entire Faster R-CNN network is established with the pre-trained method, followed by the RPN network being trained with applied training data. A proposed region created by trained RPN is then used for training the Fast R-CNN prediction system. RPN and Fast R-CNN represent a combined network, and the weight of the combined network is tuned by following the predefined procedure.

According to the pre-training CNN, the images with insulators and nests are created as training data for parameter optimization in Faster R-CNN. These insulators and nests with particular types and coordinates are highlighted in the images. Using the feature extraction system, the image features can be extracted. Afterward, a proposed region is created by the RPN with a sliding window. By comparing the training data, the weight of the fully connected layer of Fast R-CNN should be optimized.

3.3. Inception v2

Inception V2 was deployed by GoogLeNet. Inception V1 (or GoogLeNet) is the modern structural design at ILSRVRC 2014 (Szegedy et al., 2016). It has developed the minimum error at the ImageNet classifier; however, it is embedded with few effective points where the enhancement is performed to achieve better accuracy and reduce the complications of the model. Previously, Inception V1 applied convolutions like 5 × 5 which results in dimension reduction by the enlarged margin. It tends to limit the NN accuracy. A basic structure of Inception V1 is shown in Fig. 2. A reason behind NN is malicious to data loss, while the input dimension is reduced in abundance. Besides, there is also difficulty reduction while applying higher convolutions such as 5 × 5 related to 3 × 3.

It can be compiled of 1 pooling layer and seven convolutional layers, those convolution kernels through several sizes mean receptive areas of various sizes. Unlike another convolutional network, Inception V2 optimized the technique by utilizing two 3 × 3 convolutional layers rather than one 5 × 5 convolution layer that implies kernel sizes only requires 1 × 1 and 3 × 3. An Inception v2 model is depicted in Fig. 3 (Comaniciu et al., 2003). Besides, the stride is the place to 1. An enhancement is reduces further parameters and introduces further nonlinear transformations that allow the network to contain further higher learning capability. Also, Batch normalization (BN) is an effective regularization technique signified in Inception V2. Relating to the predefined method accelerates the training speed of a huge convolutional network and enhances the accuracy of classification considerably.

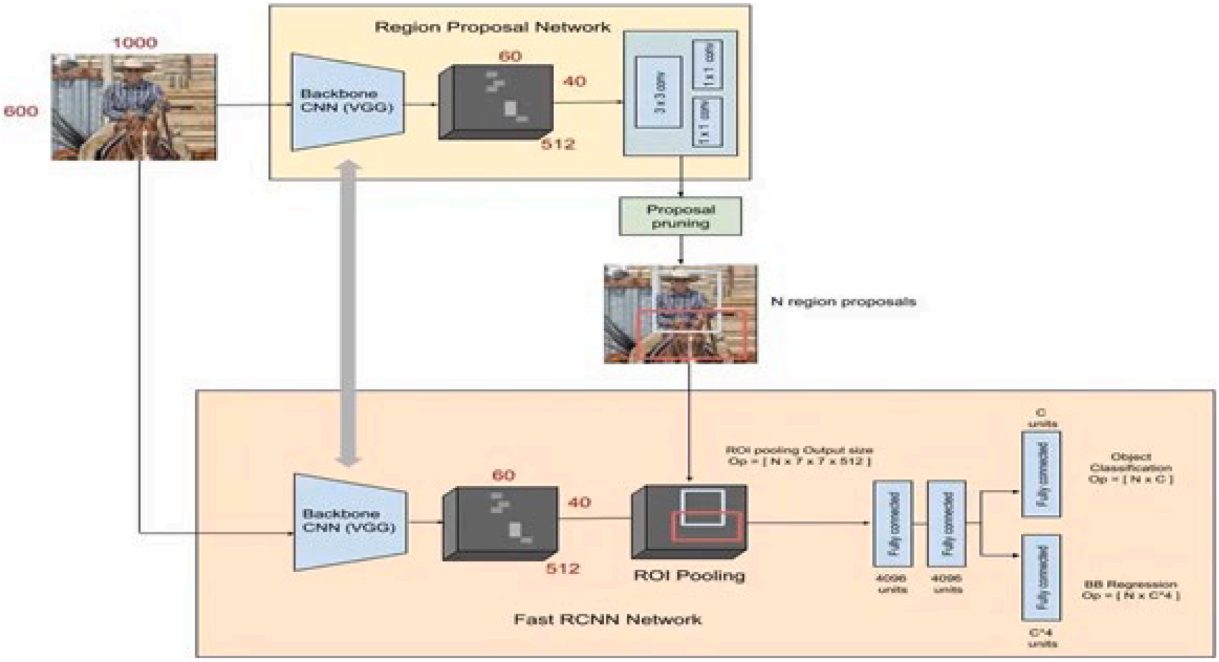


Fig. 1. Process in faster RCNN.

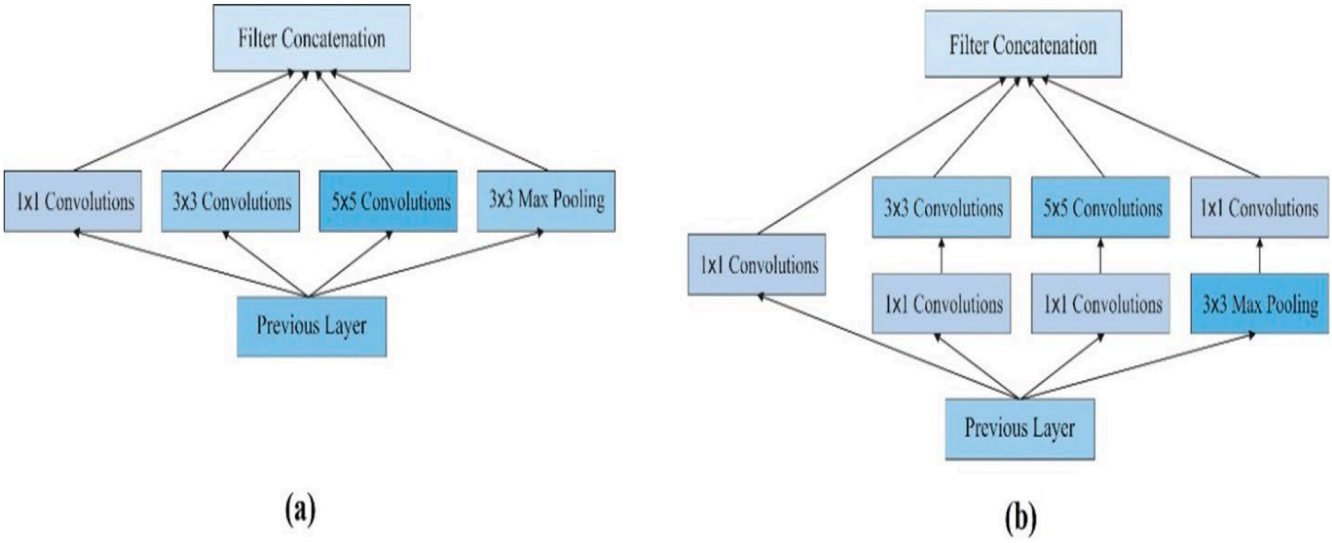


Fig. 2. a) Inception module naive version b) Inception module with dimension reductions.

The internal illustration of the testing data is in the normalized method behind executing BN into a network layer, after that a result is normalized for the normal distribution that is attributes to reduce the internal covariate shift. A weight among layers alters in batch gradient descent, and is associated with the update of activations in all hidden layers. An internal layer in the deep CNN for adapting to the data sharing is altered always that cause covariate shift. It considers that the input of a definite layer is normalization as follows:

$$\widehat{x} = \frac{x - E[x]}{\sqrt{Var[x] + \omega}} \tag{1}$$

Where x and $\widehat{x}$ indicate the input and normalization value of a definite layer correspondingly. $E[x]$ and $Var[x]$ are the expectation and difference of the input correspondingly. Also, ω signifies the offset of internal covariance. It is removed by implying BN, and a similar allocation of input is obtained in all layers behind normalized. For diminishing the impacts on all network layers behind normalization, parameters γ and β are contained in equation illustrated as follows (Comaniciu et al., 2003):

$$y_i = \gamma \widehat{x}_i + \beta \tag{2}$$

$$\frac{\partial l}{\partial \widehat{x}_i} = \frac{\partial l}{\partial y_i} \cdot \gamma \tag{3}$$

$$\frac{\partial l}{\partial \delta^2 \theta} = \sum_{i=1}^{m} \frac{\partial l}{\partial \widehat{x}_i} \cdot (x_i - \mu_\theta) \cdot \frac{-\left(\delta_\theta^2 + \omega\right)^{-3/2}}{2} \tag{4}$$

$$\frac{\partial l}{\partial \mu_\theta} = \left(\sum_{i=1}^{m} \frac{\partial l}{\partial \widehat{x}_i} \cdot \frac{-1}{\sqrt{\delta_\theta^2 + \omega}} \right) + \frac{\partial l}{\partial \delta^2 \theta} \cdot \frac{-2 \sum_{i=1}^{m} (x_i - \mu_\theta)}{m} \tag{5}$$

$$\frac{\partial l}{\partial x_i} = \frac{\partial l}{\partial \widehat{x}_i} \cdot \frac{1}{\sqrt{\delta_\theta^2 + \omega}} + \frac{\partial l}{\partial \delta^2 \theta} \cdot \frac{2(x_i - \mu_\theta)}{m} + \frac{\partial l}{\partial \mu_\theta} \cdot \frac{1}{m} \tag{6}$$

$$\frac{\partial l}{\partial \gamma} = \sum_{i=1}^{m} \frac{\partial l}{\partial \gamma_i} \cdot \widehat{x}_i \tag{7}$$

$$\frac{\partial l}{\partial \widehat{x}_i} = \frac{\partial l}{\partial \gamma_i} \cdot \gamma \tag{8}$$

$$\frac{\partial l}{\partial \beta} = \sum_{i=1}^{m} \frac{\partial l}{\partial \gamma_i} \tag{9}$$

Where l is determined as the gradient loss of back propagation. m is the size of mini-batch θ. x_i and $y_i$indicates the value of input x over the mini-batch and result behind BN model correspondingly. μ_θ and δ_θ^2 are the mean as well as variance of the mini-batch. The end result of BN network y is illustrated as follows:

$$y = \frac{\gamma x}{\sqrt{Var[x] + \omega}} + \beta - \frac{\gamma E[x]}{\sqrt{Var[x] + \omega}} \tag{10}$$

On other hand, Inception implementing BN diminishes several some many internal covariate shifts for normalizing the result in all layers. Conversely, it reduces the count of parameters and accelerates the calculating speed. The details related to the layers in Inception v2 are shown in Table 3.

4. Performance validation

Here, the wider series of experiments are performed on three dataset and the attained outcomes are examined concerning for to prediction accuracy, annotation time, Center Location Error (CLE), and Overlap Rate (OR). The information of a dataset, measures, and results analysis is defined in the upcoming sections. RS-CNN, Fast R-CNN, MDT, MPPCA and SF were applied to process the comparison task.

4.1. Dataset used

Table 4 offers the details of the 3 test dataset. Dataset 1 is defined as a multi-object tracking bird dataset (Grabner et al., 2010), which is composed of 99 frames with period time of 3 s. The second UCSDped2 (Test004) is said to be an anomalous detection dataset (Ahn & Shin, 2018), which is comprised of 180 frames and a time limit of the video is 6 s. Dataset 3 contains Two sub-files such as underwater blurred as well as underwater crowded (Zhu et al., 2018) with 2875 and 4600 frames under the time limit of 575 s.

4.2. Results analysis on dataset 1

Fig. 4 visualizes the detection of multiple objectives by the AIA-IFRCNN model on the applied dataset 1. The input image with the respected tracked outcome is illustrated and it is depicted that the bounding box identified the chick, pelican, and cloud toy objects with respective tracking accuracy.

Fig. 5 illustrates the comparative analysis of the AIA-IFRCNN model on the applied dataset 1 in terms of detection accuracy. The figure stated

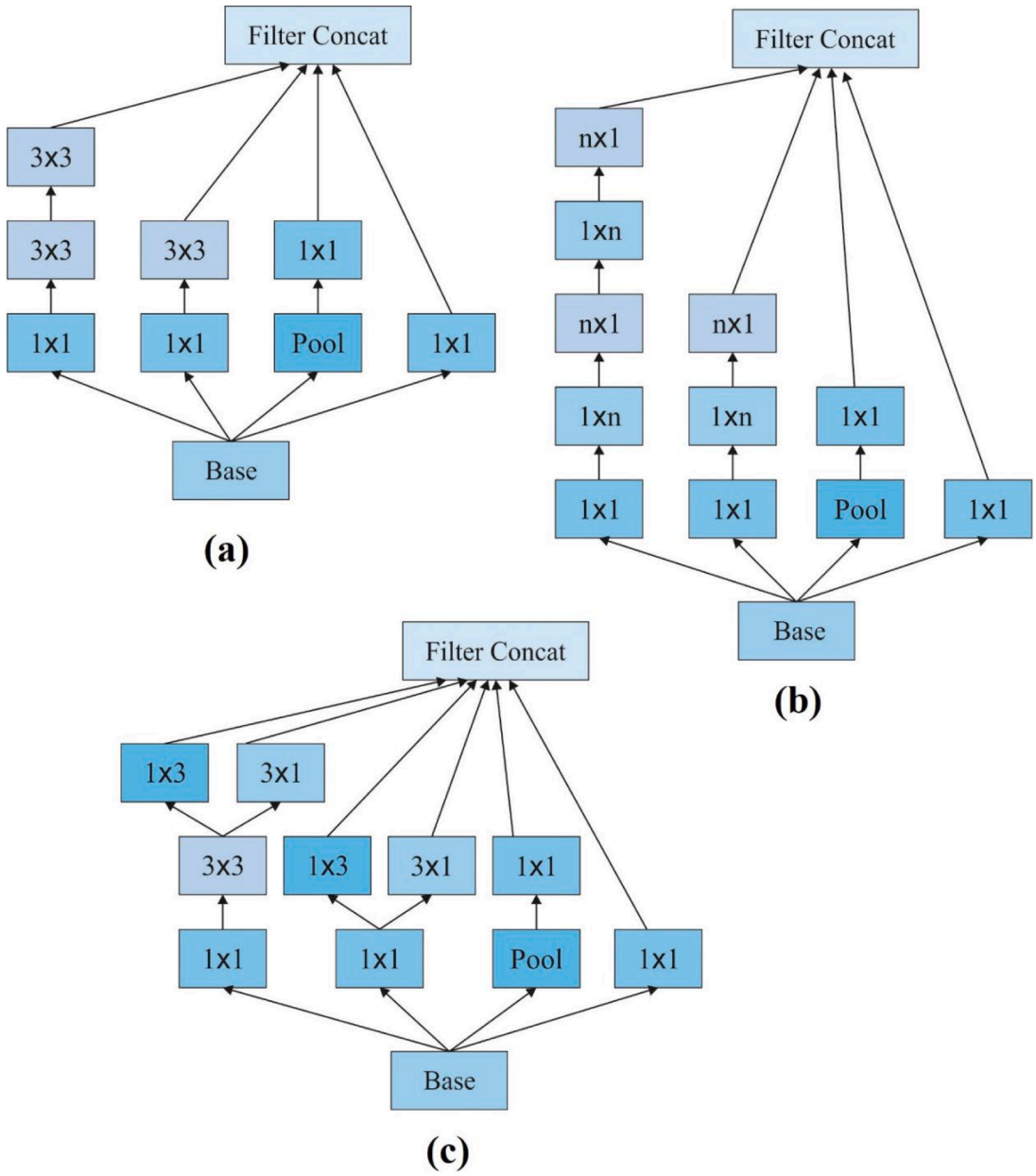


Fig. 3. a) 5 × 5 convolution is now represented as two 3 × 3 convolutions b) Two 3 × 3 convolutions, are represented as 1 × 3 and 3 × 1 in series c) inception module wider.

Table 3
Layer in inception V2.

Type	Patch Size/Stride	Input Size
Conv	3 x 3/2	299 x 299 x 3
Conv	3 x 3/1	149 x 149 x 32
Conv padded	3 x 3/1	147 x 147 x 32
Pool	3 x 3/2	147 x 147 x 64
Conv	3 x 3/1	73 x 73 x 64
Conv	3 x 3/2	71 x 71 x 80
Conv	3 x 3/1	35 x 35 x 192
3 x Inception	As in figure	35 x 35 x 288
5 x Inception	As in figure	17 x 17 x 768
2 x Inception	As in figure	8 x 8 x 1280
Pool	8 x 8	8 x 8 x 2048
Linear	logits	1 x 1 x 2048
Softmax	classifier	1 x 1 x 1000

Table 4
Dataset details.

Samples	Dataset Name	Frames	Time (s)
Dataset 1 (hanyang.ac.kr)	Bird	99	3
Dataset 2 (bib49)	UCSDped2 (Test004)	180	6
Dataset 3 (Krishnaraj et al., 2020)	Under Water (Blurred)	2875	575
	Under Water (Crowded)	4600	575

that the SF model has depicted as an ineffective tracker, which has reached the least detection accuracy over the compared methods. Simultaneously, the MPPCA model has exhibited higher accuracy than the SF model. On the other hand, the MDT model has tried to surpass the previous models. Besides, the Fast R-CNN has demonstrated manageable results with moderate detection accuracy. Following by, the RS-CNN and AIA-RFRCNN models have portrayed competitive results with high detection accuracy. However, the AIA-IFRCNN model has achieved

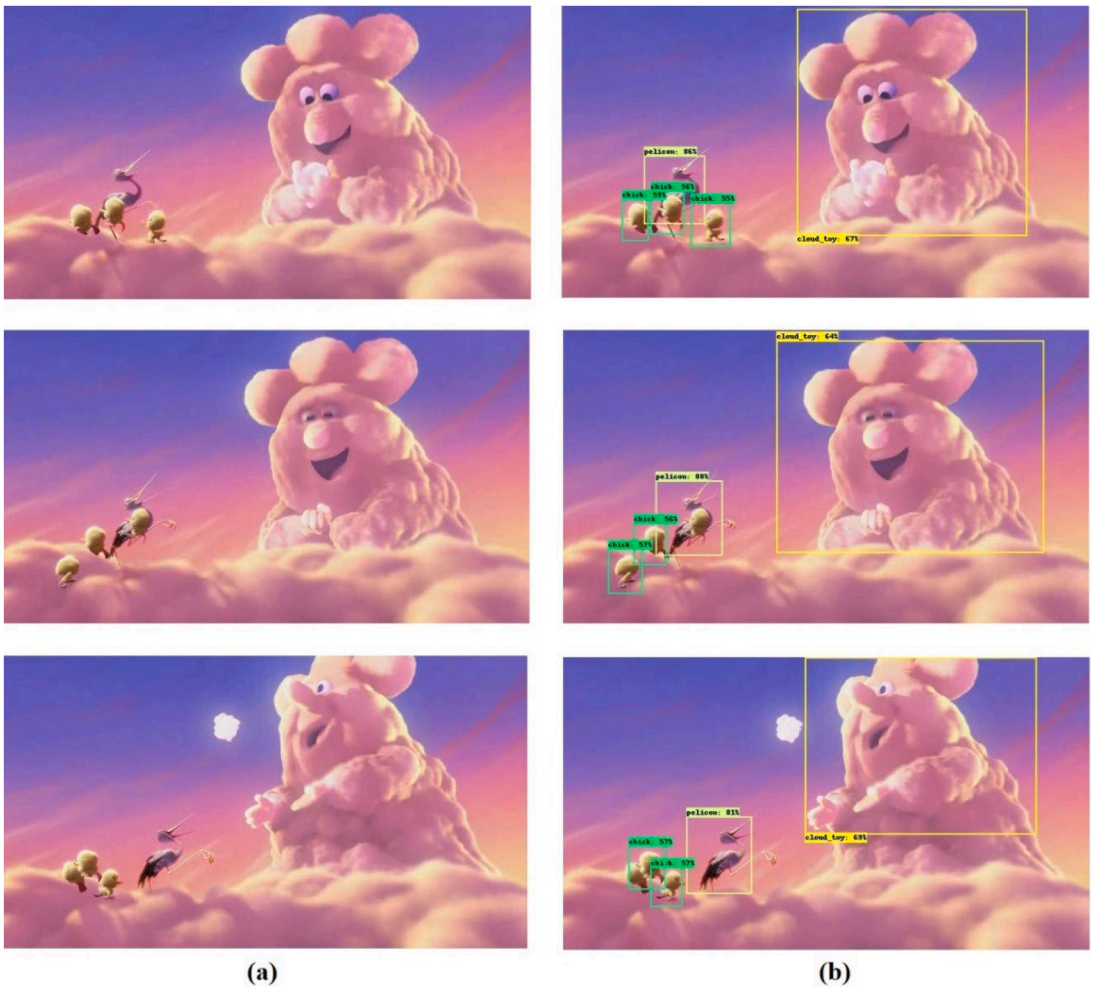


Fig. 4. Visualizing objection detection of AIA-IFRCNN for dataset 1.

superior performance by attaining maximum detection accuracy.

Fig. 6 investigates the average detection accuracy of the presented AIA-IFRCNN model with compared methods on the applied dataset 1. The figure depicted that the SF model has offered a detection accuracy of 67.01 %, which is lower than the performance attained by other methods. Followed by, the MPPCA and MDT models have achieved slightly higher detection accuracy of 73.12 % and 77.70 % respectively. Besides, the Fast R-CNN model can attain moderate detection accuracy of 86.91 % whereas the RS-CNN method offers even better results with detection accuracy of 93 %. Though the AIA-RFRCNN model has achieved a considerable detection accuracy of 94.67 %, the presented AIA-IFRCNN model has resulted in effective performance with a higher detection accuracy of 95.62 %.

4.3. Results analysis on dataset 2

Fig. 7 imagines the prediction of diverse objectives using the AIA-IFRCNN method on dataset 2. The input image with corresponding tracked results is depict that the person, car, truck, and skater objects are found by bounding box with parallel tracking accuracy.

Fig. 8 depicts the relative analysis of the AIA-IFRCNN approach on the applied dataset 2 concerning detection accuracy. The figure implied that the SF approach has showcased the worst tracker which attained lower prediction accuracy than the previous technologies. At the same time, the MPPCA scheme has represented maximum accuracy when compared to the SF model. Followed by, the MDT scheme has attempted to perform well than the existing methodologies. Then, the Fast R-CNN and RS-CNN technologies have depicted considerable outcomes with better detection accuracy. Next, the IRS-CNN and AIA-RFRCNN approaches have depicted competing results with higher detection accuracy. However, the AIA-IFRCNN technique has accomplished supreme function by reaching optimal detection accuracy.

Fig. 9 examines the average detection accuracy of the projected AIA-IFRCNN scheme with previous models on the applied dataset 2. The figure portrayed that the SF approach has provided a detection accuracy of 56.38 %, which is less than the function performed by alternate models. Besides, the MPPCA and MDT methodologies have accomplished moderate detection accuracy of 74.56 % and 81.11 % correspondingly. Followed by, the Fast R-CNN scheme is capable of reaching considerable detection accuracy of 85.10 % while acceptable results are generated by the RS-CNN method with a detection accuracy of 97.50 %. Although the IRS-CNN and AIA-RFRCNN approaches have attained a reasonable detection accuracy of 97.77 % and 98.43 %, the projected AIA-IFRCNN scheme has resulted effectual performance with a maximum detection accuracy of 98.85 %.

4.4. Results analysis on dataset 3

Fig. 10 showcases the detection of various objectives by the AIA-IFRCNN method on the applied dataset 3. The input image with the given tracked result is demonstrated and it is illustrated that the fish objects are discovered by the bounding box with corresponding tracking accuracy.

Fig. 11 depicts the competing analysis of the AIA-IFRCNN method on

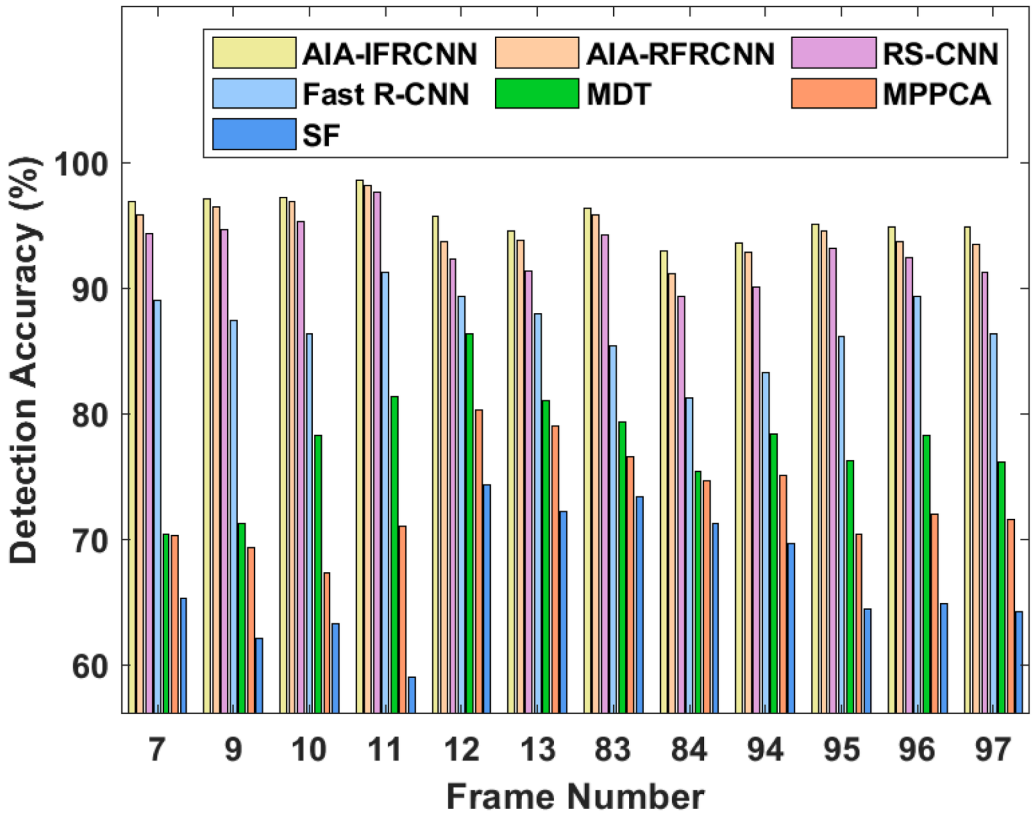


Fig. 5. Detection accuracy analysis of AIA-IFRCNN model on dataset 1.

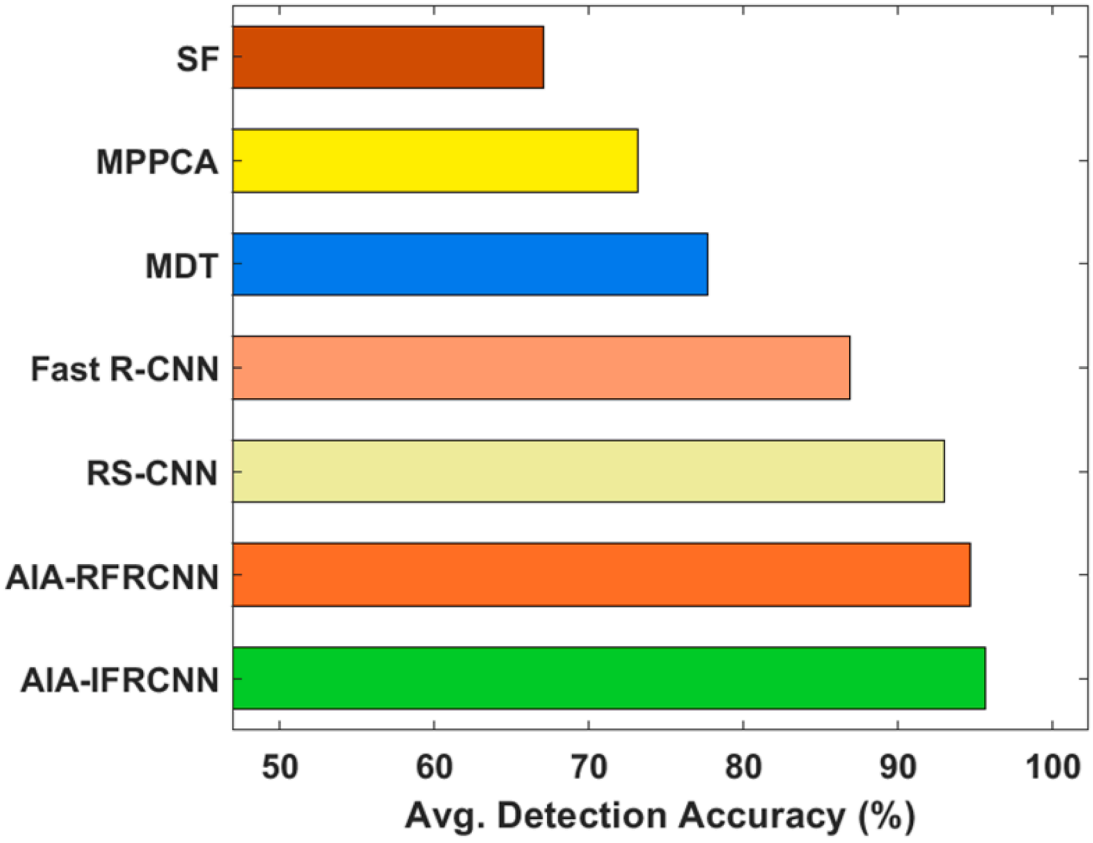


Fig. 6. Average detection accuracy analysis of AIA-IFRCNN model on dataset 1.

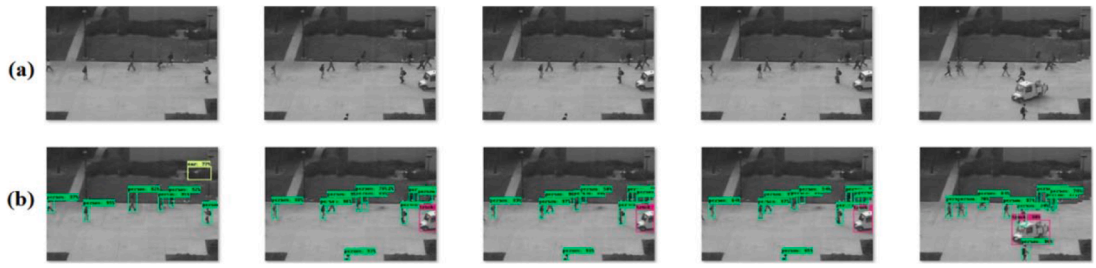


Fig. 7. Visualizing objection detection of AIA-IFRCNN for dataset 2.

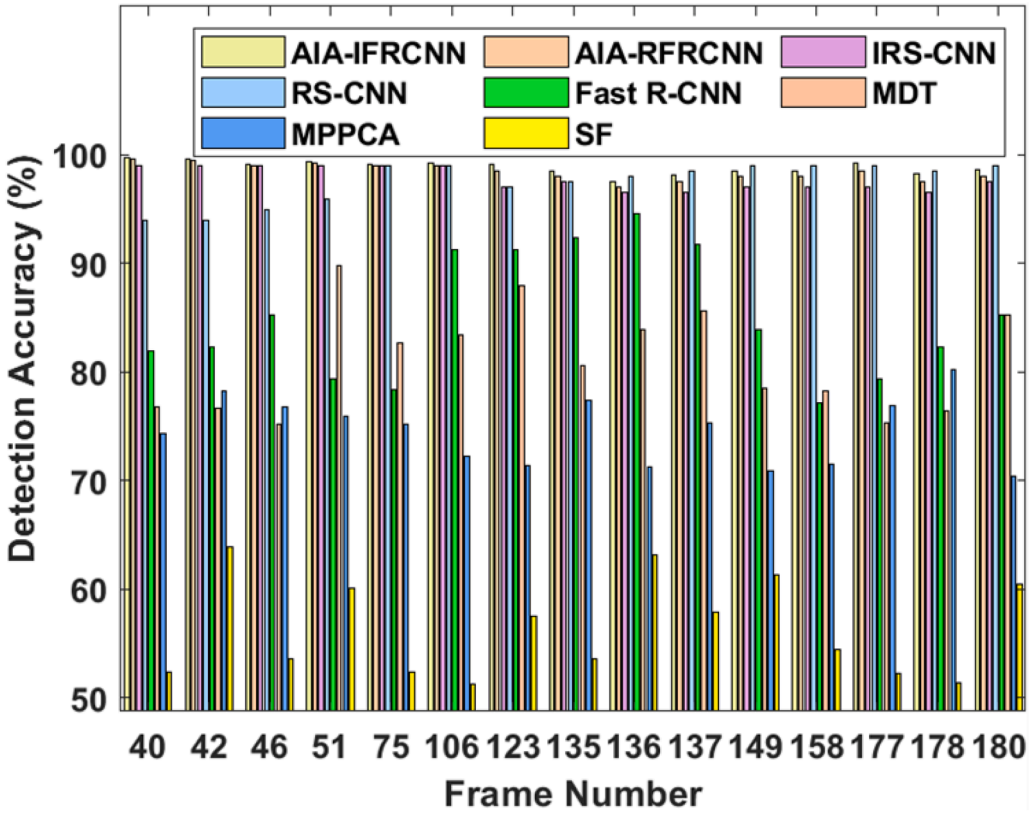


Fig. 8. Detection accuracy analysis of AIA-IFRCNN model on dataset 2.

the applied dataset 3 in light of detection accuracy. The figure depicted that the SF approach has showcased an inferior tracker that accomplished minimum detection accuracy compared to the former models. Concurrently, the MPPCA scheme has shown maximum accuracy compared to the SF model. Besides, the MDT approach has managed to outperform the traditional methods. On the other hand, the Fast R-CNN has illustrated considerable results with considerable detection accuracy. Then, the RS-CNN and AIA-RFRCNN methodologies have implied competing results with higher detection accuracy. However, the AIA-IFRCNN technology has accomplished supreme function by achieving the best detection accuracy.

Fig. 12 examines the average detection accuracy of the presented AIA-IFRCNN model with previous models on the applied dataset 3. The figure demonstrated that the SF technology has provided a detection accuracy of 66.94 %, which is minimal to the performance achieved by alternate schemes. Then, the MPPCA and MDT methodologies reached acceptable detection accuracy of 75.95 % and 82.62 % correspondingly. Followed by, this the Fast R-CNN technology has the potential to obtain better detection accuracy of 91.05 % while even better results can be generated by the RS-CNN scheme with a detection accuracy of 94.23 %. Even though the AIA-RFRCNN approach has attained a reasonable detection accuracy of 96.15 %, the projected AIA-IFRCNN scheme has provided an efficient function with a maximum detection accuracy of 97.77 %.

4.5. Analysis of average CLE

Fig. 13 investigates the performance of the AIA-IFRCNN model on the applied dataset in terms of average CLE. On analyzing the average CLE results in dataset 1, the presented AIA-IFRCNN model shows its effectiveness by attaining a minimum average CLE of 4.16. At the same time, the other methods such as AIA-RFRCNN, OMFL, CSK, FCT, DSST, CF2 and KCF models have resulted in a higher average CLE of 5.67, 7.49, 17.38, 90.30, 56.72, 38.50 and 45.57. When determining the average CLE results in dataset 2, the newly developed AIA-IFRCNN method showcases the effectiveness efficiency by obtaining the least average CLE of 5.78. Meanwhile, the alternate models like AIA-RFRCNN, OMFL, CSK, FCT, DSST, CF2, and KCF approaches have achieved maximum values of CLE of 6.89, 9.21, 19.48, 15.86, 58.31, 40.58 and 47.20. On investigating the average CLE results in dataset 3, the proposed AIA-

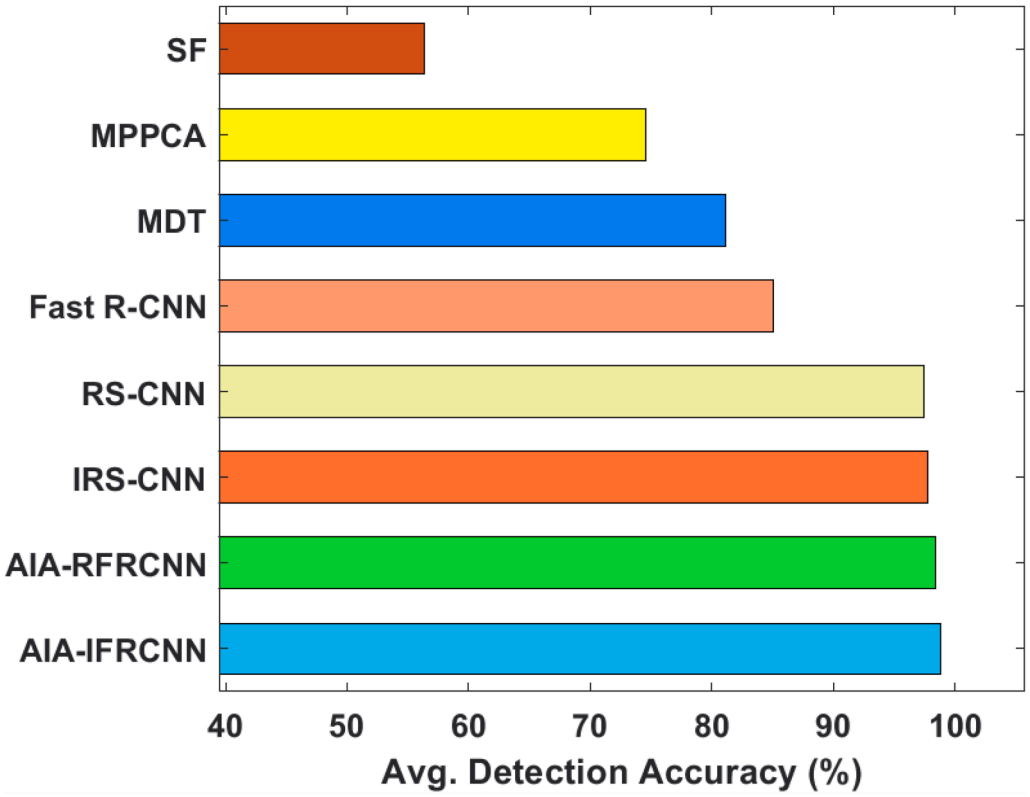


Fig. 9. Average detection accuracy analysis of AIA-IFRCNN model on dataset 2.

Fig. 10. Visualizing objection detection of AIA-IFRCNN for Dataset 3 (Crowded).

IFRCNN scheme implies supremacy by acquiring lower average CLE of 3.54. Simultaneously, the other schemes like AIA-RFRCNN, OMFL, CSK, FCT, DSST, CF2, and KCF models have generated maximum average CLE of 4.32, 12.88, 29.40, 20.79, 63.84, 48.76 and 50.42.

4.6. Analysis of overlap rate

An analysis of the overlap rate by the AIA-IFRCNN model has been made with the existing methods on the applied dataset, and the outcomes are depicted in Fig. 14. The figure showcased that the AIA-IFRCNN model has achieved superior results over the existing techniques by attaining maximum overlap rate. When determining the overlap rate in dataset 1, the projected AIA-IFRCNN model has achieved a higher overlap rate of 0.92 in contrast the AIA-RFRCNN, OMFL, CSK, FCT, DSST, CF2, and KCF models have displayed lower overlap rates of 0.89, 0.78, 0.71, 0.74, 0.52, 0.68 and 0.63 respectively. On analyzing the overlap rate in dataset 2, the newly developed AIA-IFRCNN method has reached a maximum overlap rate of 0.90 while the AIA-RFRCNN, OMFL, CSK, FCT, DSST, CF2 and KCF methodologies have showcased the least overlap rates of 0.86, 0.77, 0.68, 0.73, 0.49, 0.65 and 0.59 correspondingly. On investigating the overlap rate in dataset 3, the presented AIA-IFRCNN scheme has reached a greater overlap rate of 0.94 and the AIA-RFRCNN, OMFL, CSK, FCT, DSST, CF2, and KCF technologies have exhibited minimum overlap rates of 0.91, 0.72, 0.64, 0.69, 0.46, 0.62 and 0.51 correspondingly.

5. Conclusion

This paper has developed a novel DL based object detection and tracker model utilizing the AIA-IFRCNN model. The presented model performs the detects and tracks of objects in a sequence of frames. Initially, the DCF-CFRT model is used to annotate the objects that exist in the image. Next, Faster RCNN is applied as an object detector, which also includes the inception v2 model as the shared CNN. Finally, the softmax layer based classification process is carried out. The effectiveness of the AIA-IFRCNN method undergoes experimentation against three benchmark datasets, such as Bird (Dataset 1), UCSDped2 (Test004) (Dataset 2), Under Water (Blurred & Crowded) (Dataset 3) for

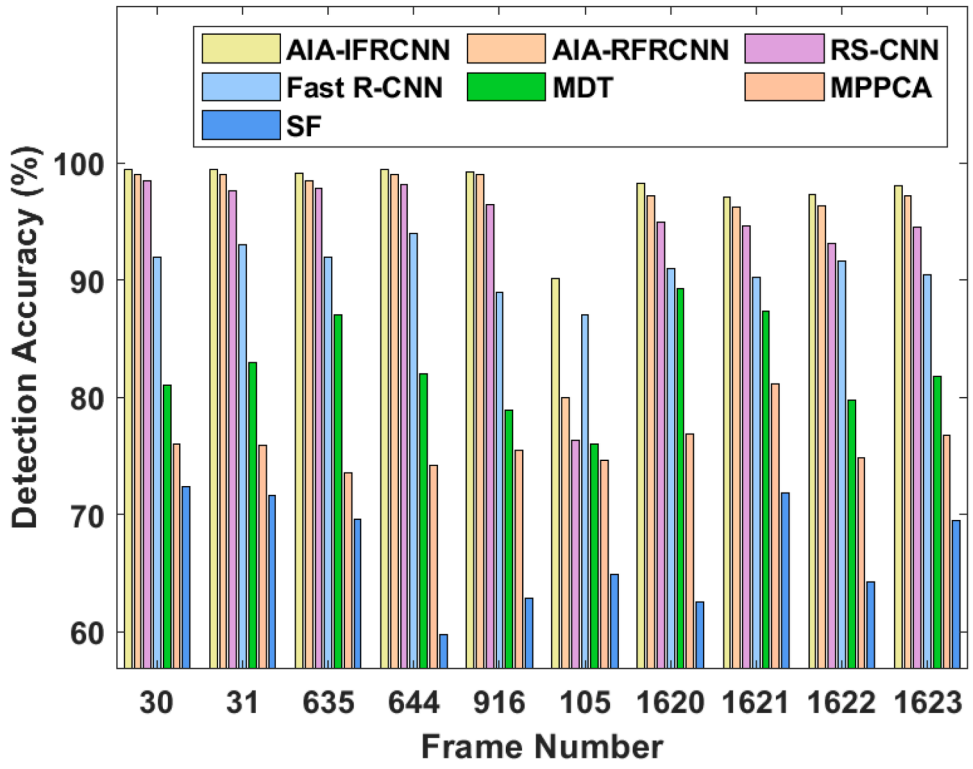


Fig. 11. Detection accuracy analysis of AIA-IFRCNN model on dataset 3.

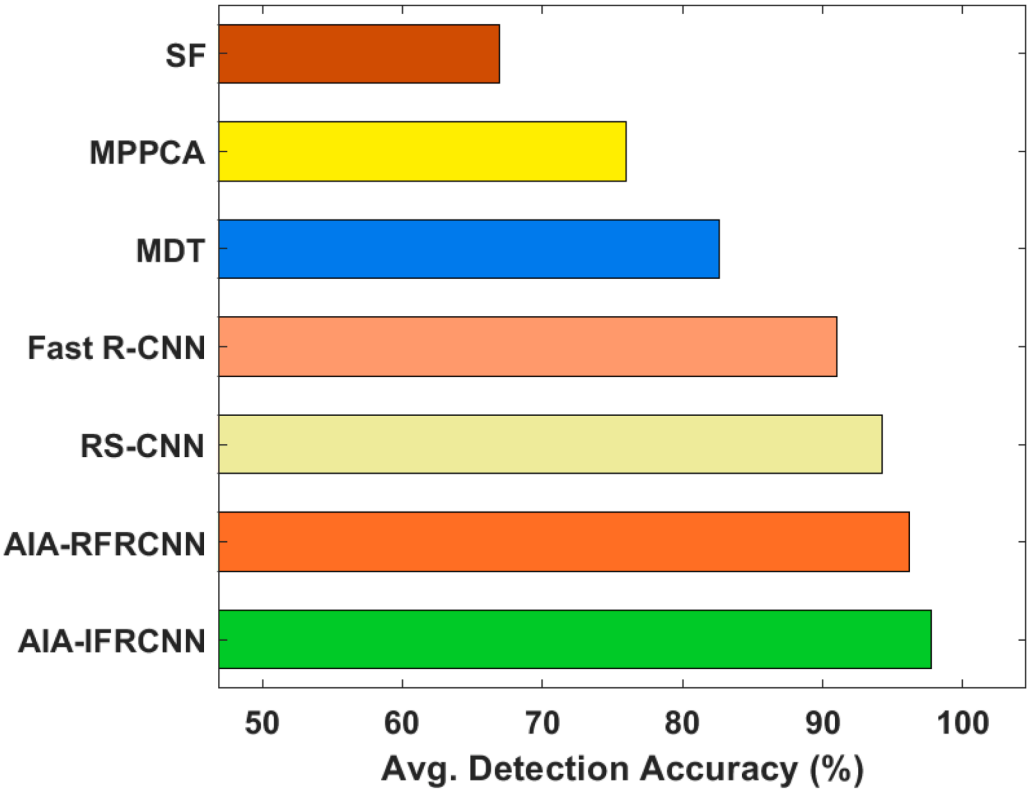


Fig. 12. Average detection accuracy analysis of AIA-IFRCNN model on dataset 3.

determining their Prediction accuracy, Annotation time, Center Location Error (CLE) and Overlap Rate (OR). The experimental outcome indicated that the AIA-IFRCNN model has outperformed the compared methods with the maximum detection accuracy of 95.62 %, 98.85 % and 97.77 % on datasets I, II, and III respectively and minimum average CLE of 4.16, 5.78 and 3.54 for the three datasets, respectively. Moreover, a higher overlap rate of 0.92, 0.90 and 0.94 was achieved for the three datasets compared to other models indicating the superiority of the

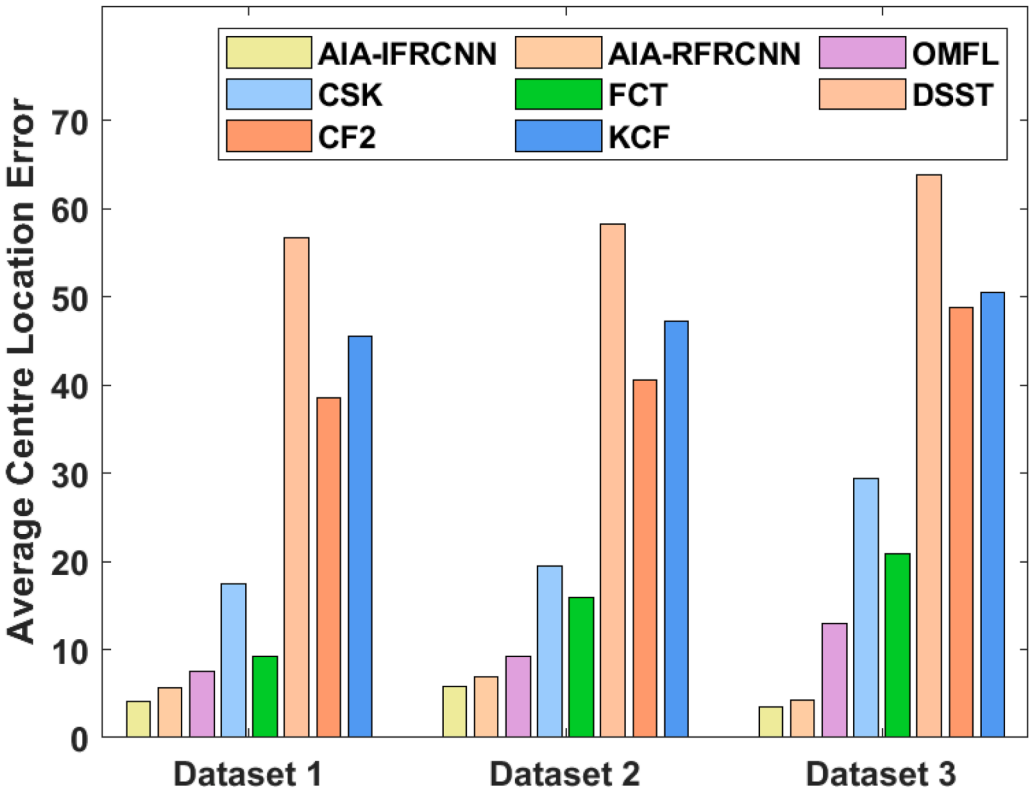


Fig. 13. Average CLE analysis of AIA-IFRCNN model.

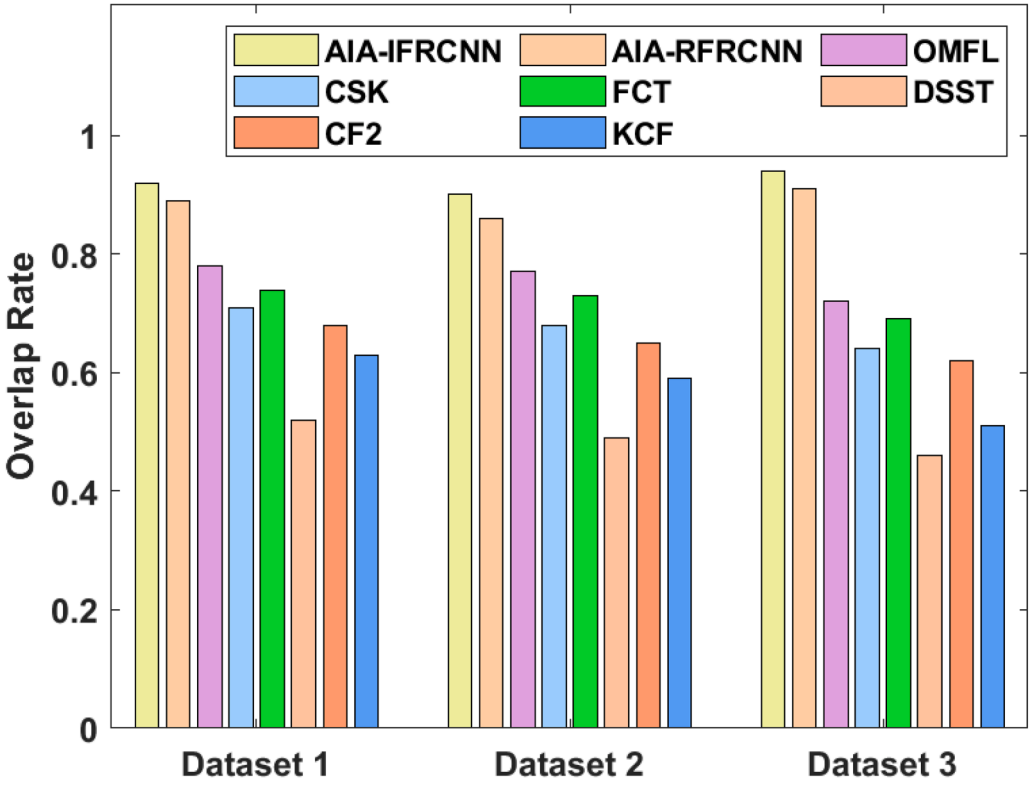


Fig. 14. Overlap rate analysis of AIA-IFRCNN model.

proposed models over other models. The experimental outcome confirmed the effective tracking performance of the projected method which can be deployed in concurrent surveillance cameras to detect and tracking abnormalities. The experimental outcome indicated that the AIA-IFRCNN model has outperformed the compared methods. Faster RCNN is a two stage detector with higher localization & recognition

accuracy but lower inference speed. In future, we intend to develop more hybrid models using single stage detectors with real time datasets. Single stage detectors provide more accuracy and higher speed. This paradigm enables researchers to the understand in the underwater object and tracking circumstances better. The research included engineering graduates with an excellent platform to learn about the better model for evaluating different underwater object detection and tracking based on deep learning-based techniques.

CRediT authorship contribution statement

K. Vijiyakumar: Conceptualization, Data curation, Investigation, Methodology, Writing – original draft. **V. Govindasamy:** Resources, Software, Supervision, Writing – review & editing. **V. Akila:** Resources, Software, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Deori, B., & Thounaojam, D. (Jul. 2014). A survey on moving object tracking in video. *International Journal on Information Theory, 3*, 31–46. https://doi.org/10.5121/ijit.2014.3304

Schmidhuber, J. (2015). Deep Learning in neural networks: An overview. *Neural Networks, 61*, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003

Kamate, S., & Yilmazer, N. (2015). Application of object detection and tracking techniques for unmanned aerial vehicles. *Procedia Computer Science, 61*, 436–441. https://doi.org/10.1016/j.procs.2015.09.183

Müller, M., Bibi, A., Giancola, S., Alsubaihi, S., & Ghanem, B. (2018). TrackingNet: A large-scale dataset and benchmark for object tracking in the wild. In *Lect. notes comput. sci. (including subser. lect. notes artif. intell. lect. notes bioinformatics), 11205* pp. 310–327). LNCS. https://doi.org/10.1007/978-3-030-01246-5_19

Henriques, J. F., Caseiro, R., Martins, P., & Batista, J. (2015). High-speed tracking with kernelized correlation filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 37*(3), 583–596. https://doi.org/10.1109/TPAMI.2014.2345390

Comaniciu, D., & Ramesh, V. (2000). Real-time tracking of non-rigid objects using mean shift 3 bhattacharyya coe cient based metric for target localization. In *, 2. Comput. Vis. Pattern Recognition, IEEE Conf. on.* (pp. 142–149).

Comaniciu, D., Ramesh, V., & Meer, P. (2003). Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 25*(5), 564–577. https://doi.org/10.1109/TPAMI.2003.1195991

Goyal, B., Dogra, A., & Sangaiah, A. K. (2022). An effective nonlocal means image denoising framework based on non-subsampled shearlet transform. *Soft Computing, 26*(16), 7893–7915. https://doi.org/10.1007/s00500-022-06845-y

Wahab, H., Mehmood, I., Ugail, H., Sangaiah, A. K., & Muhammad, K. (2023). Machine learning based small bowel video capsule endoscopy analysis: Challenges and opportunities. *Future Generation Computer Systems, 143*, 191–214. https://doi.org/10.1016/j.future.2023.01.011

Grabner, H., Matas, J., Van Gool, L., & Cattin, P. (2010). Tracking the invisible: Learning where the object might be. In *2010 IEEE computer society conference on computer vision and pattern recognition* (pp. 1285–1292). https://doi.org/10.1109/CVPR.2010.5539819

H. Ahn and I. Shin, "Study on a robust object tracking algorithm based on improved SURF method with CamShift," vol. 23, no. 1, pp. 41–48, 2018.

Held, D., Thrun, S., & Savarese, S. (2016). *Learning to track at 100 fps with deep regression networks bt - Computer Vision – eccv 2016* (pp. 749–765). Test.

Bertinetto, L., Valmadre, J., Henriques, J. F., Vedaldi, A., & Torr, P. H. S. (2016). Fully-convolutional siamese networks for object tracking. In *Lect. notes comput. sci. (including subser. lect. notes artif. intell. lect. notes bioinformatics), 9914* pp. 850–865). LNCS. https://doi.org/10.1007/978-3-319-48881-3_56

Li, B., Yan, J., Wu, W., Zhu, Z., & Hu, X. (2018). High performance visual tracking with siamese region proposal network. In *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit* (pp. 8971–8980). https://doi.org/10.1109/CVPR.2018.00935

Murugan, B. S., Elhoseny, M., Shankar, K., & Uthayakumar, J. (2019). Region-based scalable smart system for anomaly detection in pedestrian walkways. *Computers & Electrical Engineering : An International Journal, 75*, 146–160. https://doi.org/10.1016/j.compeleceng.2019.02.017

Chan, A. B., & Vasconcelos, N. (2008). Modeling, clustering, and segmenting video with mixtures of dynamic textures. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 30*(5), 909–926. https://doi.org/10.1109/TPAMI.2007.70738

Kim, J., & Grauman, K. (2009). Observe locally, infer globally: A space-time MRF for detecting abnormal activities with incremental updates. In *, 2009. 2009 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work. CVPR Work. 2009* (pp. 2921–2928). IEEE. https://doi.org/10.1109/CVPRW.2009.5206569. no. June.

Henriques, J. F., Caseiro, R., Martins, P., & Batista, J. (2012). Exploiting the circulant structure of tracking-by-detection with kernels. In *Lect. notes comput. sci. (including subser. lect. notes artif. intell. lect. notes bioinformatics), 7575* pp. 702–715). LNCS. https://doi.org/10.1007/978-3-642-33765-9_50. no. PART.

Zhang, K., Zhang, L., & Yang, M. H. (2014). Fast compressive tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 36*(10), 2002–2015. https://doi.org/10.1109/TPAMI.2014.2315808

Danelljan, M., Häger, G., Khan, F. S., & Felsberg, M. (2014). Dsst. In *BMVC 2014 - Proc. Br. Mach. Vis. Conf* (p. 2014).

Ma, C., Bin Huang, J., Yang, X., & Yang, M. H. (2019). Robust visual tracking via hierarchical convolutional features. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 41*(11), 2709–2723. https://doi.org/10.1109/TPAMI.2018.2865311

Mehran, R., Oyama, A., & Shah, M. (2009). Abnormal crowd behavior detection using social force model. In *, 2009. 2009 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work. CVPR Work. 2009* (pp. 935–942). IEEE. https://doi.org/10.1109/CVPRW.2009.5206641.

Ramachandran, A., & Sangaiah, A. K. (2021). A review on object detection in unmanned aerial vehicle surveillance. *International Journal of Cognitive Computing in Engineering, 2*, 215–228. https://doi.org/10.1016/j.ijcce.2021.11.005

Mittal, P., Sharma, A., & Singh, R. (2022). A simulated dataset in aerial images using simulink for object detection and recognition. *International Journal of Cognitive Computing in Engineering, 3*, 144–151. https://doi.org/10.1016/j.ijcce.2022.07.001

Chen, L., et al. (2020). Underwater object detection using invert multi-class adaboost with deep learning. In *Proc. Int. Jt. Conf. Neural Networks*. https://doi.org/10.1109/IJCNN48605.2020.9207506

Yang, H., Liu, P., Hu, Y., & Fu, J. (2021). Research on underwater object recognition based on YOLOv3. *Microsystem Technologies : Sensors, Actuators, Systems Integration, 27*(4), 1837–1844. https://doi.org/10.1007/s00542-019-04694-8

Liu, C., et al. (2022). A new dataset, Poisson GAN and Aquanet for underwater object grabbing. *IEEE Transactions on Circuits and Systems for Video Technology, 32*(5), 2831–2844. https://doi.org/10.1109/TCSVT.2021.3100059

Hu, K., Lu, F., Lu, M., Deng, Z., & Liu, Y. (2020). A marine object detection algorithm based on ssd and feature enhancement. *Complexity, 2020*. https://doi.org/10.1155/2020/5476142

R. Gavrilescu, C. Fo, C. Zet, and D. Cotovanu, "Faster R-CNN : An approach to real-time object detection," pp. 165–168, 2018.

Sommer, L., Schmidt, N., Schumann, A., & Beyerer, J. (2018). Search area reduction fast-RCNN for fast vehicle detection in large aerial imagery. In *Proc. - Int. Conf. Image Process. ICIP* (pp. 3054–3058). https://doi.org/10.1109/ICIP.2018.8451189

Irisa, S., & France, F. V. (2018). *BURIED object detection from b-scan ground penetrating radar data using faster-rcnn minh-tan pham* (pp. 6808–6811). S ' ebastien Lef ` evre Universit '.

Manana, M., Tu, C., & Owolawi, P. A. (2018). Preprocessed faster RCNN for vehicle detection. In *2018 International conference on intelligent and innovative computing applications (ICONIC)* (pp. 1–4). https://doi.org/10.1109/ICONIC.2018.8601243

Wang, B., Xu, J., Li, J., Hu, C., & Pan, J. S. (2017). Scene text recognition algorithm based on faster RCNN. In *2017 First international conference on electronics instrumentation & information systems (EIIS)* (pp. 1–4). https://doi.org/10.1109/EIIS.2017.8298720

Ning, C., Zhou, H., Song, Y., & Tang, J. (2017). Inception Single Shot MultiBox Detector for object detection. In *2017 IEEE International conference on multimedia & expo workshops (ICMEW)* (pp. 549–554). https://doi.org/10.1109/ICMEW.2017.8026312

Chen, K. H., Shou, T. D., Li, J. K. H., & Tsai, C. M. (2018). Vehicles detection on expressway via deep learning: single shot multibox object detector. In *, 2. 2018 International conference on machine learning and cybernetics (ICMLC)* (pp. 467–473). https://doi.org/10.1109/ICMLC.2018.8526958

Sangari, M. S., Thangaraj, K., Vanitha, U., Srikanth, N., Sathyamoorthy, J., & Renu, K. (2023). Deep learning-based object detection in underwater communications system. In *2023 2nd Int. Conf. Electr. Electron. Inf. Commun. Technol. ICEEICT 2023* (pp. 1–6). https://doi.org/10.1109/ICEEICT56924.2023.10157072

Aruna, S. K., Deepa, N., & Devi, T. (2023). Underwater fish identification in real-time using convolutional neural network. In *Proc. 7th Int. Conf. Intell. Comput. Control Syst. ICICCS 2023* (pp. 586–591). https://doi.org/10.1109/ICICCS56967.2023.10142531

Serdà, R. F., & Burguera, A. (2023). *Using deep neural networks to detect and track fish in underwater video sequences* (pp. 1–7). Limerick, Ocean. Limerick: Ocean. https://doi.org/10.1109/OCEANSLimerick52467.2023.10244295. *2023 -2023*.

Sun, B., Zhang, W., Su, Z., & Wang, H. (2023). Real-time underwater target tracking using PP-YOLO and cloud computing. In *2023 6th Int. Symp. Auton. Syst. ISAS 2023* (pp. 1–6). https://doi.org/10.1109/ISAS59543.2023.10164579

Li, B., & Lima, D. (2021). Facial expression recognition via ResNet-50. *International Journal of Cognitive Computing in Engineering., 2*, 57–64. https://doi.org/10.1016/j.ijcce.2021.02.002

Liang, X., Zhang, J., Zhuo, L., Li, Y., & Tian, Q. (2020). Small object detection in unmanned aerial vehicle images using feature fusion and scaling-based single shot detector with spatial context analysis. *IEEE Transactions on Circuits and Systems for Video Technology, 30*(6), 1758–1770. https://doi.org/10.1109/TCSVT.2019.2905881

Chandan, G., Jain, A., Jain, H., & Mohana. (2018). Real time object detection and tracking using deep learning and OpenCV. In *2018 International conference on inventive research in computing applications (ICIRCA)* (pp. 1305–1308). https://doi.org/10.1109/ICIRCA.2018.8597266

Krishnan, V., Vaiyapuri, G., & Govindasamy, A. (2022). Hybridization of deep convolutional neural network for underwater object detection and tracking model. *Microprocessors And Microsystems, 94*, Article 104628. https://doi.org/10.1016/j.micpro.2022.104628

Lei, X., & Sui, Z. (2019). Intelligent fault detection of high voltage line based on the Faster R-CNN. *Measurement Journal of the International Measurement Confederation, 138*, 379–385. https://doi.org/10.1016/j.measurement.2019.01.072

"faster-r-cnn-for-object-detection-a-technical-summary-474c5b857b46 @ towardsdatascience.com." .

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. *Proceedings of the IEEE conference on computer vision and pattern Recognition*, 2818–2826. https://doi.org/10.1109/CVPR.2016.308, 2016-Decem.

Zhu, X., Zhu, M., & Ren, H. (2018). Method of plant leaf recognition based on improved deep convolutional neural network. *Cognitive Systems Research, 52*, 223–233. https://doi.org/10.1016/j.cogsys.2018.06.008

"datasets @ cvlab.hanyang.ac.kr." .

"dataset @ www.svcl.ucsd.edu." .

Krishnaraj, N., Elhoseny, M., Thenmozhi, M., Selim, M. M., & Shankar, K. (2020). Deep learning model for real-time image compression in Internet of Underwater Things (IoUT). *Journal of Real-Time Image Processing, 17*(6), 2097–2111. https://doi.org/10.1007/s11554-019-00879-6