DOI: 10.24996/ijs.2025.66.12.38

ISSN: 0067-2904

Detecting, Tracking, and Calculating the Speed of Colored Balls Using Deep Learning

Heba Kh. Abbas[1], Haidar J. Mohamad[2]*, Ali A. Alzuky[2], Batool Redha Abd[1], Anwar H. Al-Saleh[3]

[1]Department of Physics, College of Science for Women, University of Baghdad, Baghdad, Iraq
[2]Department of Physics, College of Science, Mustansiriyah University, Baghdad, Iraq
[3]Department of Computer Science, College of Science, Mustansiriyah University, Baghdad, Iraq

Received: 28/6/2024 Accepted: 2/12/2024 Published: 30/12/2025

Abstract

Developing tracking in various applications motivates researchers to explore this field. In this work, an intelligent system is suggested to automatically detect moving colored balls in real time and calculate the speed and direction of these balls based on the deep learning algorithm RCNN. The data is compared with the Alex-net algorithm because it is a standard method. The proposed algorithm is one of the machine learning algorithms based on the principle of training and learning, which relies on the sequential classifier. The proposed system consists of a phone camera, colored balls, and different environmental lighting (changing from one to eight lights). There are two luxmeters used to measure the intensity of light. Four parameters are measured to evaluate the performance of algorithms and system setup: accuracy, average time, detection ratio, and speed. The best class and training were selected and approved for detecting the blue and green balls. This proposed algorithm can be used to detect any moving object. Results showed a high quality of ball detection and tracking with almost 100% accuracy.

Keywords: Detecting moving objects, RCNN, Alex-net algorithm, artificial intelligence, colored balls speed.

اكتشاف، تتبع، وحساب سرعة الكرات الملونة باستعمال التعلم العميق

هبة خضير عباس[1]، حيدر جواد محمد[2]، علي عبد داود[2]، بتول رضا عبد[1]، أنوار حسن الصالح[3]

[1]قسم الفيزياء، كلية العلوم للبنات، جامعة بغداد، بغداد، العراق
[2]قسم الفيزياء، كلية العلوم، الجامعة المستنصرية، بغداد، العراق
[3]قسم علوم الحاسوب، كلية العلوم، الجامعة المستنصرية، بغداد، العراق

الخلاصة

إن تطوير التتبع في التطبيقات المختلفة يحفز الباحثين على استكشاف هذا المجال. في هذا العمل، تم اقتراح نظام ذكي لاكتشاف الكرات الملونة المتحركة تلقائيًا في الوقت الفعلي وحساب سرعة واتجاه هذه الكرات بناءً على خوارزمية التعلم العميق RCNN. تتم مقارنة البيانات مع خوارزمية Alex-net لأنها طريقة قياسية.

* Email: Haidar.mohamad@uomustansiriyah.edu.iq

الخوارزمية المقترحة هي إحدى خوارزميات التعلم الآلي التي تعتمد على مبدأ التدريب والتعلم والتي تعتمد على المصنف المتسلسل. يتكون النظام المقترح من كاميرا هاتف، وكرات ملونة، وإضاءة بيئية مختلفة (تتراوح من مصباح واحد إلى ثمانية مصابيح). هناك نوعان من اللوكسميتر يستخدمان لقياس شدة الضوء. هناك أربع معلمات يتم قياسها لتقييم أداء الخوارزميات وإعداد النظام: الدقة ومتوسط الوقت ونسبة الكشف والسرعة. تم اختيار أفضل فئة وأفضل تدريب واعتمادها في الكشف عن الكرات الزرقاء والخضراء. يمكن استخدام هذه الخوارزمية المقترحة للكشف عن أي جسم متحرك. أظهرت النتائج جودة عالية في اكتشاف الكرة وتتبعها بدقة تصل إلى 100% تقريبًا .

1. Introduction

Moving object detection is the process of segmenting targets of interest. Identifying a moving target is the basic step in the process of classifying and tracking a moving object [1]. The main goal of motion activity object detection and tracking is forward detection of the moving target(s) either at each video frame or at the start of the moving target in the video [2, 3]. Detecting moving objects is one of the fundamental tasks in digital processing because it has a prominent role in many aspects of real-world applications [4].

Automate video analysis has three main orders: detection, tracking, and recognition [5]. In the first step, it is important to identify and segment interesting objects in video [6]. Tracking object from frame to frame becomes easy, and their trace paths can be analyzed to learn about the object's behavior [7]. Thus, object detection plays a crucial role in specific applications [8]. For example, video surveillance is used in places that require sensitive security measures, such as banks, shopping malls, and highways [9]. Detecting moving objects has always proven to be a difficult task due to many factors, such as background dynamics, illumination variations, poor classification of shadow as an object, and blurring and smoothing issues [10]. Much research work has been conducted to deal with the above factors where smart grid and security monitoring systems are used to develop smart cities [11]. These techniques are used in observation systems for higher speeds detection in modern countries [12]. This process depends on image processing algorithms and deep learning [13].

2. Related work

Many previous studies focused on identifying and detecting moving objects and targets. In 2021, Gustafsson and Persson [14], conducted a study to evaluate three different object detection techniques for detecting small and fast-moving objects in the sport of Padel. The goal was to identify which technique performs best under various conditions affecting detection performance. The three techniques employed different approaches, including background histogram calculation, HSV masking with edge detection, and deep neural network frameworks with the COCO dataset. Outdoor video footage was used to test the techniques, revealing that Canny edge detection showed promising high detection accuracy. YOLO demonstrated a good ability to detect many objects with high confidence, offering reliable and accurate detection. The study addressed the challenge of introducing detection techniques for small, fast-moving objects in sports, such as Padel. It concluded that both Canny and YOLO hold promise for this application.

In 2023 , Jaemin Cho *et al*. [15] , used selective attention networks to detect objects in an industrial environment in real time. They used CCTV cameras to detect workers moving in a factory while doing their job. The challenge to this approach lied in a person who is covered by the vehicle or other objects. The recognition distance was 0.6 m, and the maximum speed was 2.6 m/s. They recommended improving object detection by creating 3D maps for the sites to detect any object in case of being covered by one side.

In 2023, Fareeha Mumtaz *et al*. [16], presented a comprehensive review of R-CNN and YOLO techniques applied in moving objects detection. The review paper focused on traffic jams where the vehicles must be recognized while it is moving. The Lucas-Kanade method is used to calculate motion vectors in each following frame. They show that the presented technique has advantages in detecting vehicles with low cost and high accuracy, although the cars are moving very slowly.

In 2023[17], Li *et al*. [17], proposed an efficient pipeline for detecting moving objects, focusing on dynamic object detection in video streams from traffic monitoring cameras. In their pipeline, they utilized image eccentricity analysis as a pre-processing step to quickly generate segmentation maps of moving objects. These maps were used to mask the original images, isolating only the moving objects. Sparse images were then inputted into an object detection model built with a sparse convolution backbone network, significantly reducing computational costs. Quantitative experiments demonstrated that the proposed pipeline achieved up to a 50% inference speedup with negligible detection accuracy loss in images obtained from traffic monitoring cameras.

In 2023, Tchao *et al*. [18], discussed Phase 1 of the AnxEpiVR pilot study, aimed at exploring scenarios that provoke epilepsy/seizure-specific interictal anxiety and laying the groundwork for designing VR exposure therapy (ET) scenarios for people with epilepsy (PwE). Analysis revealed anxiety-provoking scenes categorized under themes like location, social setting, situational, activity, physiological, and previous seizure. Public settings, social situations, and factors like potential danger and specific triggers were identified as common anxiety triggers. Recommendations were made for creating customizable VR-ET scenarios incorporating different anxiety-related factors. Subsequent phases will involve creating VR-ET hierarchies and evaluating their feasibility and effectiveness.

This work presents a proposed system to track a moving black, green, and blue ball using a smartphone (iPhone 14 pro max) camera. The lightness change affects the detection quality. Therefore, eight lamps are used to test the lightness effect with RCNN algorithm. The results parameters like detection, accuracy, average time, and speed are compared with Alex-net algorithm results.

3. Theoretical Principles

Region convolutional neural network (R-CNN) is a type of deep learning method focused on object detection in AI applications [19]. It is one of the advanced models used for object detection by merging convolutional neural networks and region-based approaches [20]. R-CNN algorithm involves separate testing images into many regions or sub-regions, which named region proposals [21]. These regions are responsible for generating objects according to the sub-images. Edge Boxes are used to generate regional proposals in RCNN [22]. Features are extracted by CNN in the time of generating the region proposals, which is around 2,000 regions approximately [23]. The regions are resized to 16 pixels in the warped frame. Alex-Net used CNN to fine-tune large datasets like ImageNet for feature representation. CNN has a high-dimensional features vector for the tested region proposal. These features are sent to a separate machine-learning classifier to classify objects [24, 25]. Support Vector Machines (SVMs) are used for the classification step and generate a class for the region proposal if it has an object or not in the training process. There are two types of training output: positive and negative samples, which refer to containing an instance of the class or not[26, 27]. Non-maximum suppression is the final object detection step to ensure that only the confident and non-overlapping bounding boxes are retained [28]. Disadvantages of R-CNN are computational complexity, slow inference, and overlapping region proposals [29].

Alex-Net is a deep learning model that won in a 2012 competition by training on 1.2 million images from the ILSVRC database with 1,000 classes. Its architecture consists of approximately 60 million parameters [30]. The model comprises 5 convolutional layers, 2 normalization layers, 3 max-pooling layers, 3 fully connected layers, and one SoftMax layer for output. Dropout regularization was implemented to mitigate overfitting, and the ReLU activation function was employed for each convolutional layer [31]. Quality and performance measures were relied upon in deep learning and data analysis, as the study included several efficient measures in evaluating the quality of model performance. Four scales were used. The first measure is accuracy, which was calculated based on four factors [32]. Table 1 shows how to determine these parameters [33, 34].

• true positive (TP) refers to the correct presence of the object.
• true negative (TN) refers to the correct absence of any object.
• false positive (FP) refers to the false presence of the object.
• false negative (FN) refers to the incorrect absence of any object.

Table 1: Determine quality metrics parameters.

Actual classification	1	1	0	0
Predicated classification	0	1	0	1
result	F_N	T_P	T_N	F_P

Accuracy is a metric for evaluating the performance of classification models. It is the ratio of correct predictions to the total number of predictions. For binary classification, equation (1) accuracy can also be calculated in terms of positives and negatives as [35]:

$$Accuracy\ (Acc) = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \tag{1}$$

Detection is determined in each video frame (30 f/s). If the ball is detected, then the ratio is the ball frame over the total number of frames. The detection rate equation is [36]:

$$detection\ rate\ = \frac{T_P}{T_P + F_N} \tag{2}$$

Real-time Detection is programmed within MATLAB software to identify and locate objects of interest in real-time video sequences. The time is extracted as data within the code and presented accordingly [37].

Speed is determined by calculating the displacement between two reference points on the moving object's path. The difference was measured as: $dX = X_{i+1} - X_i$and $dy = y_{j+1} - y_j$ [22]:

$$dr = \sqrt{dX^2 + dy^2}$$

(3)

$$speed = \frac{dr}{dt} = \frac{dr}{\frac{1}{30}} = dr \times 30 \tag{4}$$

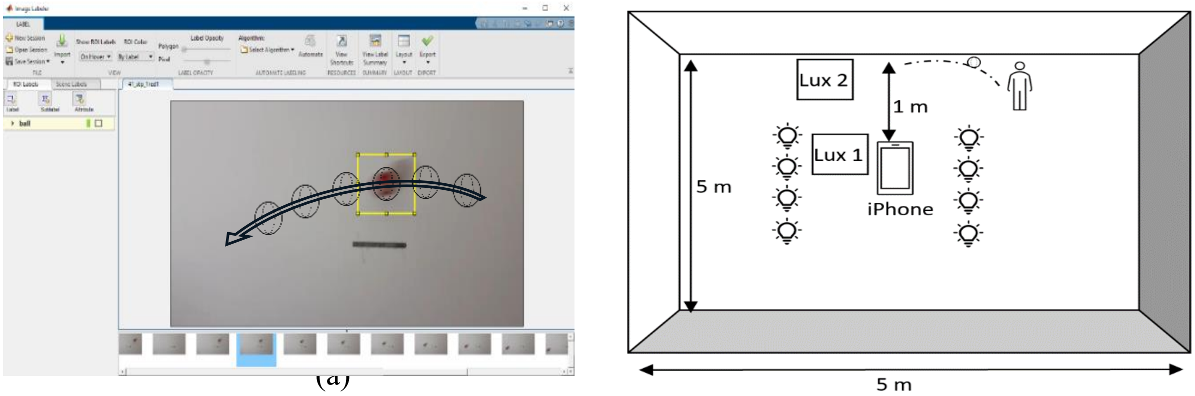
The speed was calculated in an image using the relationship:

$$speed\ = \sqrt{dX^2 + dy^2}\ \times 30\ \ \frac{pixel}{s} \tag{5}$$

The actual speed was calculated based on the scale factor:

$$scale\ factor = \frac{speed\ (\frac{m}{s})}{speed\ (\frac{pixel}{s})} \tag{6}$$

Therefore, if we need to know the speed in m/s, then we multiply 0.2419 m/pixel by speed (pixel/s). Figure 1 shows the real red ball with a yellow box from the algorithm and the estimated path with direction.



(a)

Figure 1: Ball path behavior in (a) video stream with selecting box in yellow and (b) diagram of the suggested system

4. Experimental setup

An iPhone 14 Pro Max mobile camera was used to record ten videos in different lighting (1-8 lights). Five of these videos were used for training purposes with only five blocks of lighting because we believe it is the ideal lighting, and three of them were used for testing with all lighting conditions. Two luxmeters were used to record the intensity of the environment, as shown in Table 2. These luxmeters are located next to the phone camera and next to the wall, as shown in Figure 2. In the RCNN algorithm, the layer numbers (iz) were 32, 64, and 128, while the Training times (period) were 5, 10, 20, and 50. In this case, we had twelve models generated from the five videos to use for object tracking and apply it to the three test videos with all lighting conditions. For each lighting condition (one lamp, two lamps, etc.) three videos were recorded and shown in each figure. There is a slight difference between the data of each video, which leads to a slight error due to lighting conditions. There are two main algorithms used within this work. The first algorithm is Training Data and Detection, and the second algorithm is Testing Model and Calculate Metrix.

Training Data and Detection

Input: number of layers (iz) and the number of training times(epoch), and specify a specific threshold for accuracy to stop training(acth)

- *Load number of layers iz=32, 64,128 and number of training times epoch =5, 10,20,50 and determine loss threshold =0.2, threshold for accuracy to stop training* acth=95;
- *Store the training file and the model with the same name based on the time in terms of hours and seconds*
- *Load positive folder and label file (ball. mat)*
- *Create three layers each layer contains*
- *count the number of classes since the target is a ball and the background represents two classes using fullyConnectedLayer(m)*
- *Options for training deep learning neural network*
- *Specifying a condition that the iteration is more than 300 and the accuracy is more than 95, the training stops.*
- *begin the training process as instructed based on postv, layers, option*

  [dete, info] = trainRCNNObjectDetector(pstv2, layers, options);
- To obtain the model, store it and store information about the model where det represent model contain loss and accuracy
- Finally, it stores the model, layers, options and information
- After make testing images in this model

Testing Model and Calculate Metrix

Start algorithm

- *Select Folder and select all images from 1 to ni , read images using code*

  [bbox, score, label] = detect (dete, I);
- *I f there is a target, detects it and puts box, score, label on it and its category*
- *After detection, the target is deducted, or a yellow square is placed on each detected target. This means that the score is greater than 9 else no detected*
- *For object detect calculate center x and center y because object moving then calculate average of score(accuracy)*
- *Calculate extraction time, detection time and total time then calculate average time in second*
- *Draw the path of the moving target where nx , ny= represent object path*
- *Drawing the score and calculating its average*
- *Calculate object speed during*

Xc= center x

Yc==center y

*dx= deff(xc*scf/100) in m/s*

*dy=deff(yc*scf/100) in m/s*

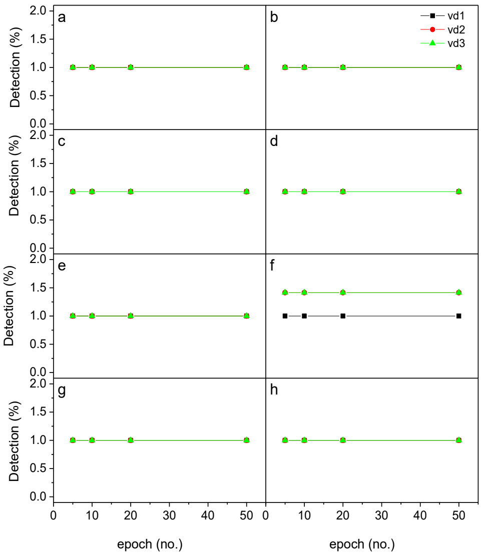
v= $sqrt(dx^2 + dy^2) * 30$

va= avage (v)

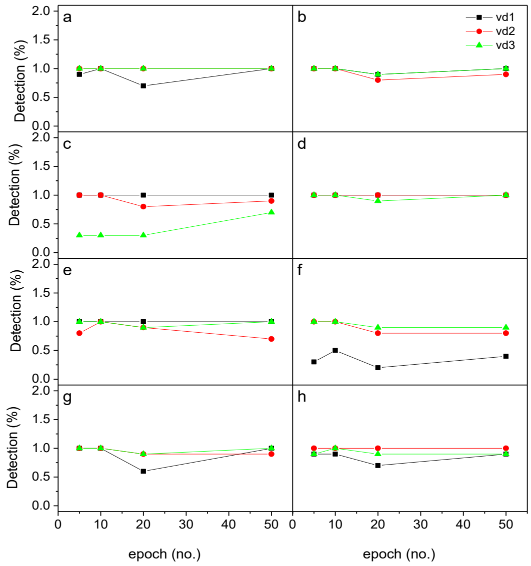
There are four parameters excluded from RCNN compared to the Alex-net method, which are detection ratio, detection accuracy, average object detection time, and object speed. For clarity, the number of layers (iz) was set to 32 in the Alex-net method. Figure 3 shows the detection percentage of different lights with different epochs, comparing Alex-Net and RCNN at different levels.

Table 2: The light intensity as distributed in the setup with the number of lights.

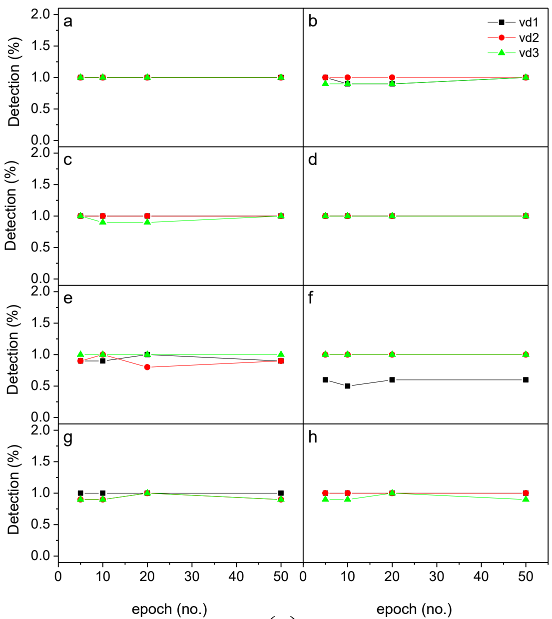
Number of lamps	0	1	2	3	4	5	6	7	8
Lux1 (Next to the camera)	2	32	63	101	127	198	203	253	317
Lux2 (Next to the wall)	2	16	41	63	81	127	198	205	227



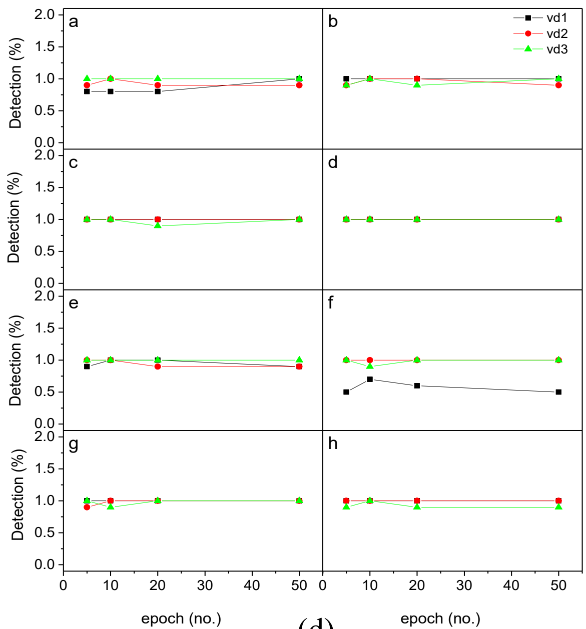
(a)



(b)



(c)



(d)

Figure 2: The relationship between epoch and object detection using (a) Alexa Net (b) RCNN iz=32, (c) RCNN iz=64, and (d) RCNN iz=128.

The accuracy was calculated based on Eq. (1), which depends on the number of correct predictions on the total number of predictions as in Figure 3.

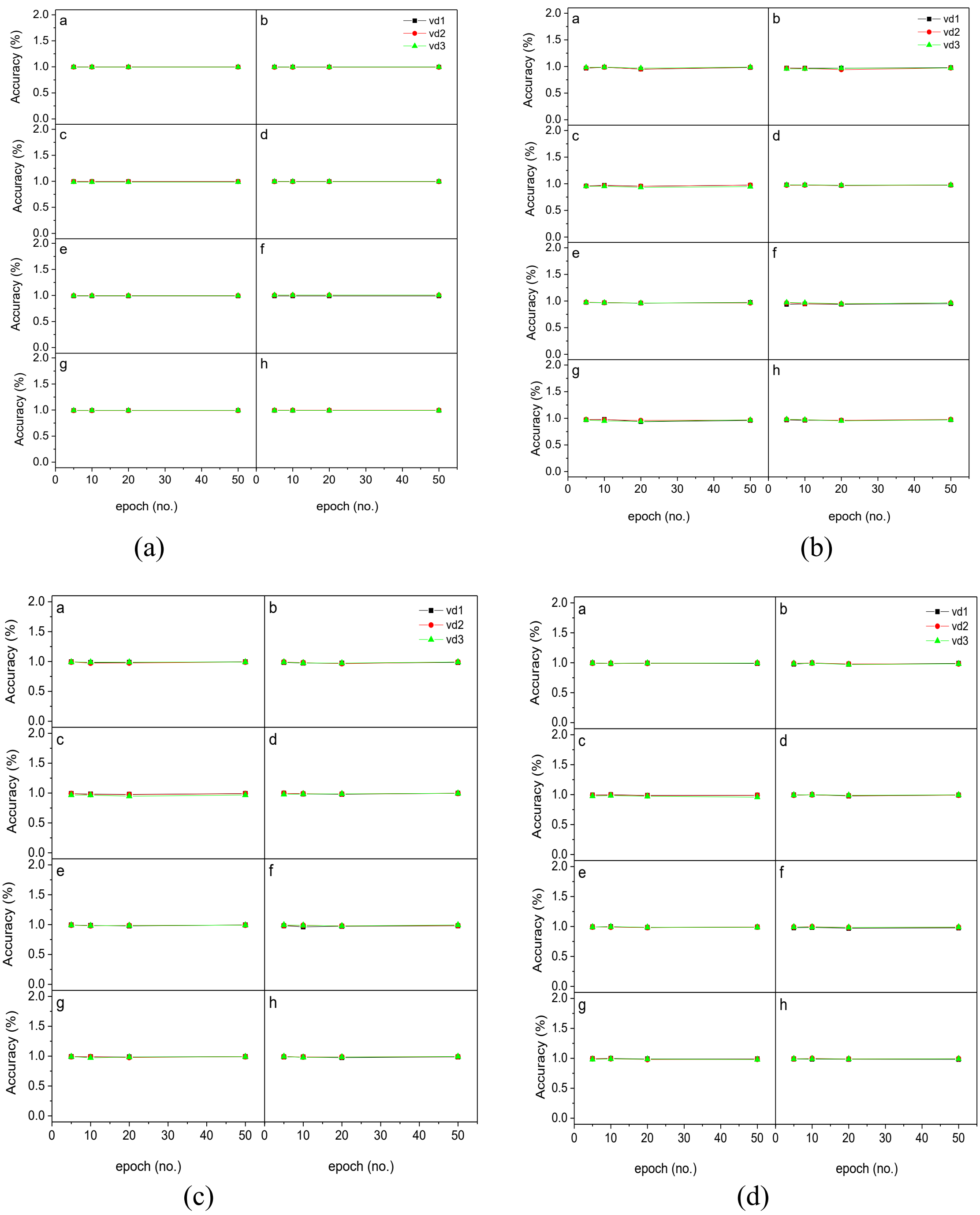
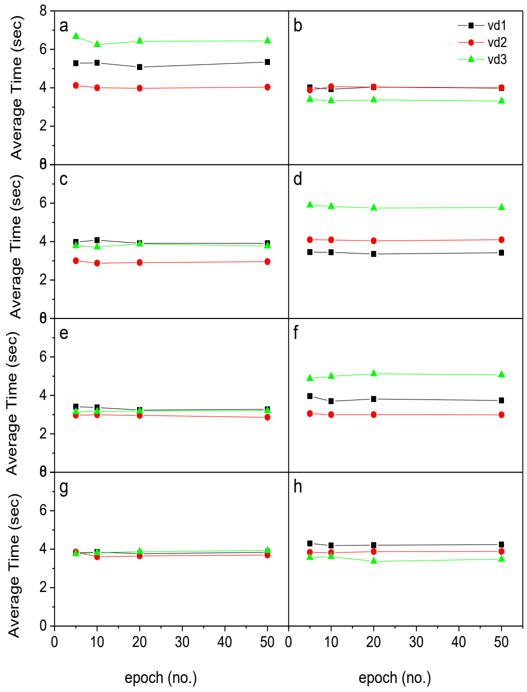
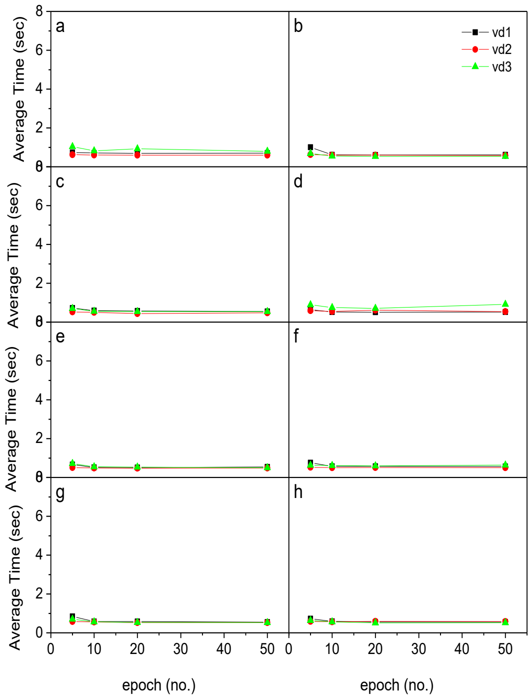


Figure 3: The relationship between epoch and Accuracy using (a) Alexa Net (b) RCNN iz=32, (c) RCNN iz=64, and (d) RCNN iz=128.

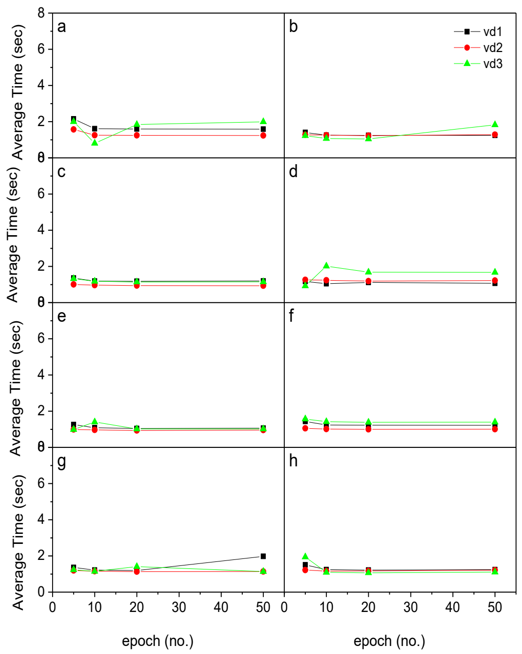
The total time rate (average time) was calculated, which represents the sum of the rate of detection time (t1) and the rate of deduction time(t2), as shown in Figure 4.



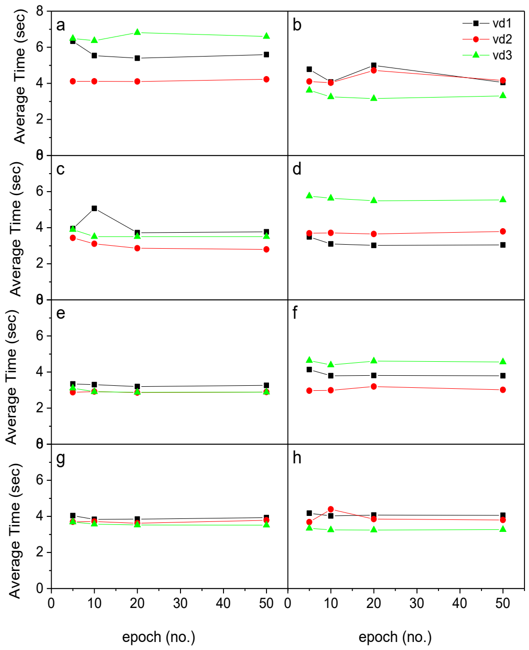
(a)



(b)



(c)



(d)

Figure 4: The relationship between epoch and Average Time using (a) Alexa Net (b) RCNN iz=32, (c) RCNN iz=64, and (d) RCNN iz=128.

The speed of the moving object was calculated based on the scale factor as explained in the third chapter, as in Figure 5, where the speed results for the moving object are shown.

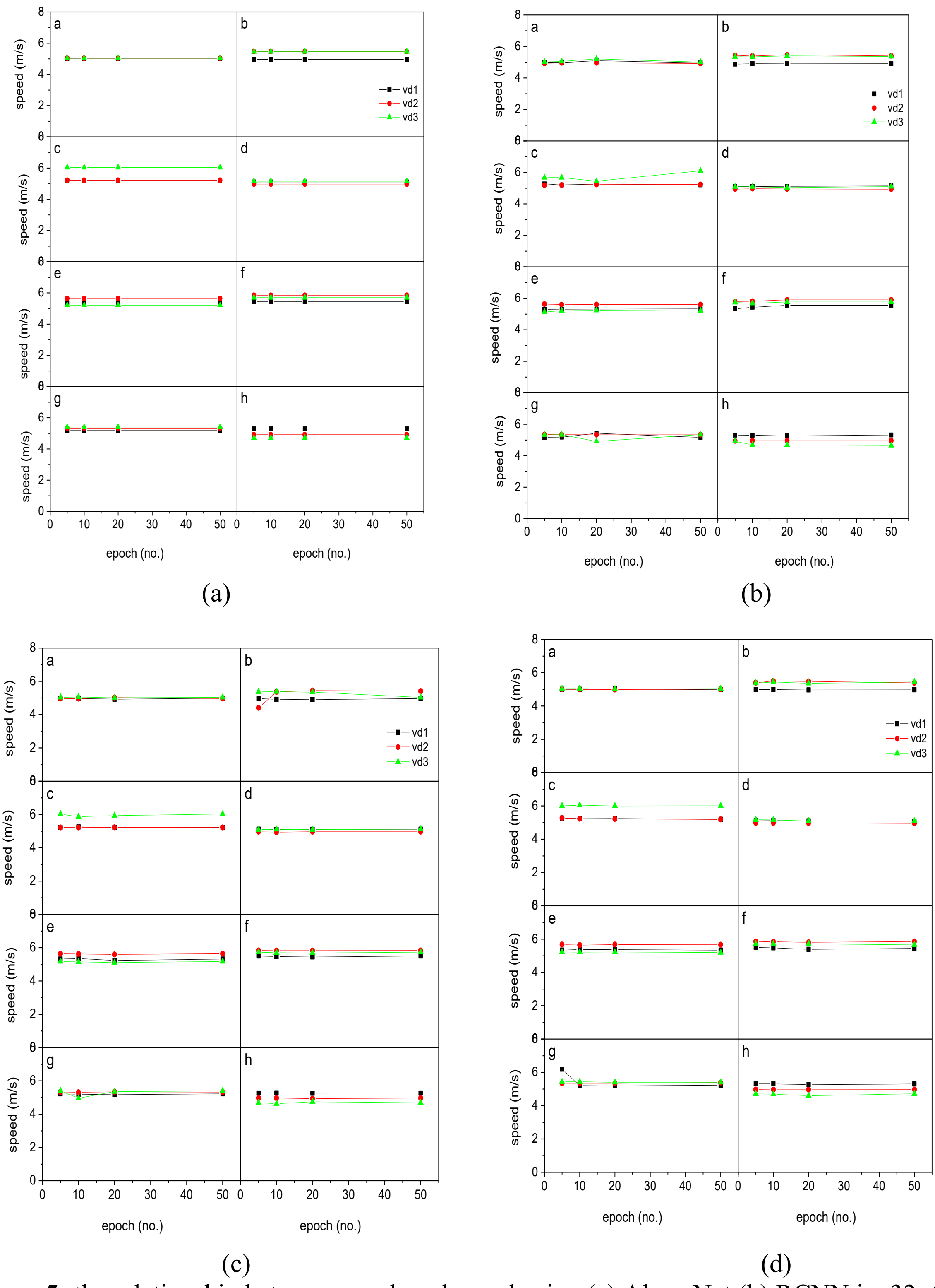
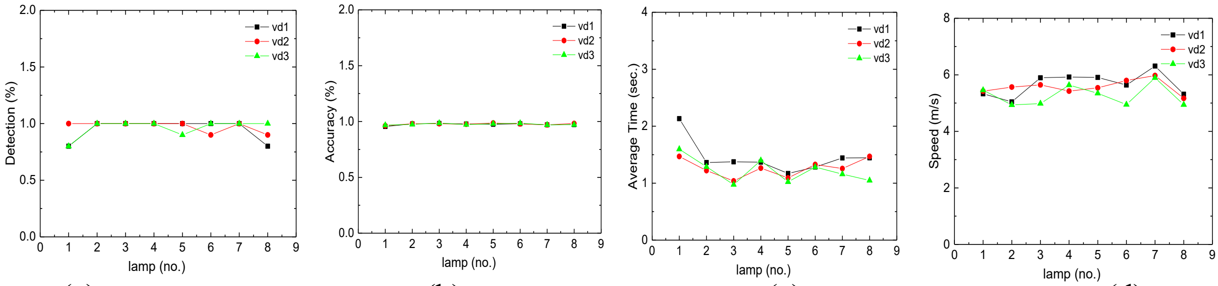


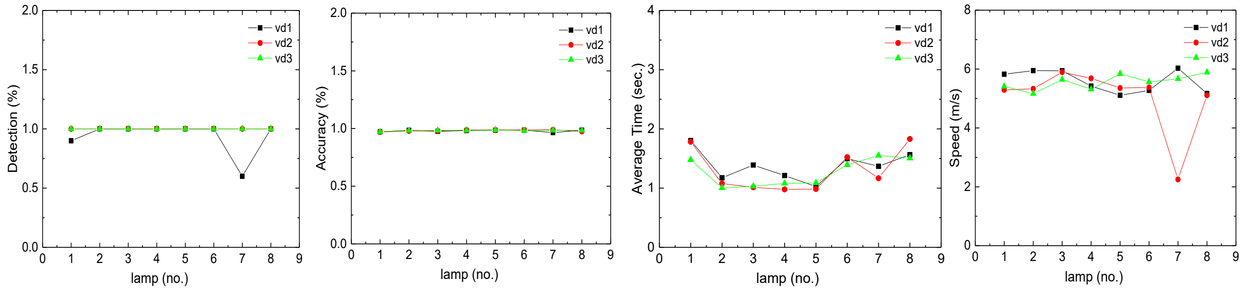
Figure 5: the relationship between epoch and speed using (a) Alexa Net (b) RCNN iz=32, (c) RCNN iz=64, and (d) RCNN iz=128.

By detecting the black ball, work was reworked on the green ball, relying on layer iz=64 and training epoch =50, because they gave the best results in the black ball. The four approved standards were calculated by working as in Figure 6 and for the blue ball as in Figure 7.



(a) (b) (c) (d)

Figure 6: Relationship between parameters (a) object detection, (b) Accuracy, (c) total average time, and (d) speed with lamp number using iz 64, epoch 50



(a) (b) (c) (d)

Figure 7: Relationship between parameter (a) object detection, (b) Accuracy, (c) total average time and (d) speed with lamp number using iz 64, epoch 50

5. Conclusion

In this study, a proposed optical system was presented to detect and measure the speed of moving colored balls based on their movement path using deep learning techniques in different lighting conditions. R-CNN method was used for this purpose and compared to the Alex-net method. Motion vectors and ball trajectories are determined from successive frames of motion data. The detection rate, accuracy, overall time rate (detection and detection), and speed of the moving object are determined.

Epoch and iz play important parameters through object detection. The change in lighting conditions does not affect the detection rate and accuracy. This leads to the conviction that the proposed algorithm is powerful enough to detect any object under different conditions with very high accuracy and quickly. The accuracy and detection rate have high values for this algorithm to detect moving objects in an average time of about 1 second and an object speed of about 6 m/s. The proposed setup is useful for checking and tracking any moving object.

Reference

[1] H. Taher and A. Tuaimah, "Detection and Classification of Moving Objects," *Educational Philosophy and Theory,* vol. 10, pp. 41-51, 02/10 2020, doi: 10.32792/utq.jceps.10.02.05.

[2] D.Venkata, S. Prasad, Ch.Niharika, R. Vishwaja, and R.Srinidhi, "Object Detection and Tracking Using Image Processing," 2020.

[3] Y.-Y. Hou, S.-Y. Chiou, and M.-H. Lin, "Real-time detection and tracking for moving objects based on computer vision method," *2017 2nd International Conference on Control and Robotics Engineering (ICCRE),* pp. 213-217, 2017.

[4] K. N. V. Khasim, G. Anilkumar, G. Vamshi, S. k. Ch, and M. Nikhil Yadav, "Moving Object Detection and Speed Estimation by Digital Image Processing," *Journal of Physics: Conference Series,* vol. 2325, no. 1, p. 012035, 2022/08/01 2022, doi: 10.1088/1742-6596/2325/1/012035.

[5] N. Funde, P. N. Paranjape, K. Ram, P. Magde, and M. M. Dhabu, "Object Detection and Tracking Approaches for Video Surveillance Over Camera Network," *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS),* pp. 1171-1176, 2019.

[6] A. Ilioudi, A. Dabiri, B. J. Wolf, and B. D. Schutter, "Deep Learning for Object Detection and Segmentation in Videos: Toward an Integration With Domain Knowledge," *IEEE Access,* vol. 10, pp. 34562-34576, 2022, doi: 10.1109/ACCESS.2022.3162827.

[7] Y. Zhang, X. Chen, J. Li, C. Wang, C. Xia, and J. Li, "Semantic Object Segmentation in Tagged Videos via Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 40, pp. 1741-1754, 2018.

[8] S. Guefrachi *et al.*, "Deep learning based DeepFake video detection," *2023 International Conference on Smart Computing and Application (ICSCA),* pp. 1-8, 2023.

[9] S. Shaikh, K. Saeed, and N. Chaki, *Moving Object Detection Using Background Subtraction*. 2014.

[10] M. Kaur and P. Maheshwari, *Building smart cities applications using IoT and cloud-based architectures*. 2016, pp. 1-5.

[11] S. Myeong, J. Park, and M. Lee, "Research Models and Methodologies on the Smart City: A Systematic Literature Review," *Sustainability*, vol. 14, no. 3, doi: 10.3390/su14031687.

[12] M. A. Mohammed *et al.*, "Machine Learning Algorithm for Developing Classroom Attendance Management System Based on Haar Cascade Frontal Face," in *2021 IEEE Symposium on Industrial Electronics & Applications (ISIEA)*, 10-11 July 2021, pp. 1-6, doi: 10.1109/ISIEA51897.2021.9509990.

[13] C. Peña, C. Rodriguez, and I. A. Romero, "Processing of Images Based on Machine Learning to Avoid Unauthorized Entry," in *2022 14th International Conference on Computational Intelligence and Communication Networks (CICN)*, 4-6 Dec. 2022, pp. 783-788, doi: 10.1109/CICN56167.2022.10008350.

[14] S. Gustafsson and A. Persson, "Detecting small and fast objects using image processing techniques : A project study within sport analysis," Independent thesis Basic level (degree of Bachelor) Student thesis, 2021. [Online]. Available: http://urn.kb.se/resolve?urn=urn:nbn:se:hj:diva-54343

[15] J. Cho and K. Kim, "Detection of moving objects in multi-complex environments using selective attention networks (SANet)," *Automation in Construction,* vol. 155, p. 105066, 2023/11/01/ 2023, doi: doi.org/10.1016/j.autcon.2023.105066.

[16] F. Mumtaz, H. Ashraf, and N. Jhanjhi, *Moving Object Detection Using Deep Learning Algorithms*. 2023.

[17] X. Li, R. Nabati, K. Singh, E. Corona, V. Metsis, and A. Parchami, "EMOD: Efficient Moving Object Detection via Image Eccentricity Analysis and Sparse Neural Networks," *2023 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW),* pp. 51-59, 2023.

[18] D. Tchao, S. Lewis-Fung, H. Gray, S. Pardini, L. Harris, and L. Appel, "Describing Epilepsy-Related Anxiety to Inform the Design of a Virtual Reality Exposure Therapy: Results from Phase 1 of the AnxEpiVR Clinical Trial," *Epilepsy & Behavior Reports,* vol. 21, p. 100588, 01/16 2023, doi: 10.1016/j.ebr.2023.100588.

[19] N. M. Hashem, H. K. Abbas, and H. J. Mohamad, "Detection and Recognition of Car Plates in Parking Lots at Baghdad University," *Iraqi Journal of Science,* vol. 64, no. 2, pp. 1018-1029, 2023, doi: doi.org/10.24996/ijs.2023.64.2.43.

[20] H. K. Abbas, A. H. Al-Saleh, H. J. Mohamad, and A. A. D. Al-Zuky, "New algorithms to Enhanced Fused Images from Auto-Focus Images," *Baghdad Science Journal,* 2020.

[21] R. Awad, A. A. D. Al-Zuky, A. H. Al-Saleh, and H. J. Mohamad, "Enhance Video Film using Retnix method," *Journal of Physics: Conference Series,* vol. 1003, 2018.

[22] H. K. Abbas, H. J. Mohamad, A. H. Al-Saleh, and A. A. Al-Zuky, "Modelling vision angles of optical camera zoom using image processing algorithm," *IOP Conference Series: Materials*

Science and Engineering, vol. 571, no. 1, p. 012117, 2019/07/01 2019, doi: 10.1088/1757-899X/571/1/012117.

[23] A. S. Derea, H. K. Abbas, A. H. Al-Saleh, and H. J. Mohamad, "Adopting Texture Features to Detect and Recognize Brain Tumors in Magnetic Resonance Images," *IOP Conference Series: Materials Science and Engineering,* vol. 757, no. 1, p. 012029, 2020/03/01 2020, doi: 10.1088/1757-899X/757/1/012029.

[24] N. H. Rasham, H. K. Abbas, A. Razaq, and H. J. Mohamad, "Simulation of Speckle Noise Using Image Processing Techniques," *Computer Networks and Inventive Communication Technologies,* 2021.

[25] M. k. Jawad, H. G. Daway, H. J. Mohamad, and E. G. Daway, "Lightness Enhancement by Fuzzy Logic Depending on Power Membership Function," *Journal of Physics: Conference Series,* vol. 1999, 2021.

[26] C. Zhan, X.-h. Duan, S. Xu, S. Zheng, and L. Min, "An Improved Moving Object Detection Algorithm Based on Frame Difference and Edge Detection," *Fourth International Conference on Image and Graphics (ICIG 2007),* pp. 519-523, 2007.

[27] H. Weiming, T. Tieniu, W. Liang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews),* vol. 34, no. 3, pp. 334-352, 2004, doi: 10.1109/TSMCC.2004.829274.

[28] R. Kaur and S. Singh, "A comprehensive review of object detection with deep learning," *Digital Signal Processing,* vol. 132, p. 103812, 2023/01/01/ 2023, doi: doi.org/10.1016/j.dsp.2022.103812.

[29] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data,* vol. 8, no. 1, p. 53, 2021/03/31 2021, doi: 10.1186/s40537-021-00444-8.

[30] D. Sudha and J. Priyadarshini, "An intelligent multiple vehicle detection and tracking using modified vibe algorithm and deep learning algorithm," *Soft Computing,* vol. 24, no. 22, pp. 17417-17429, 2020/11/01 2020, doi: 10.1007/s00500-020-05042-z.

[31] B. Zhang, L. Zhou, and J. Zhang, "A methodology for obtaining spatiotemporal information of the vehicles on bridges based on computer vision," *Computer-Aided Civil and Infrastructure Engineering,* vol. 34, no. 6, pp. 471-487, 2019/06/01 2019, doi: doi.org/10.1111/mice.12434.

[32] L. Liu, "Moving Object Detection Technology of Line Dancing Based on Machine Vision," *Mobile Information Systems,* vol. 2021, pp. 1-9, 04/27 2021, doi: 10.1155/2021/9995980.

[33] Q. Ji, C. Dai, C. Hou, and X. Li, "Real-time embedded object detection and tracking system in Zynq SoC," *EURASIP Journal on Image and Video Processing,* vol. 2021, no. 1, p. 21, 2021/06/16 2021, doi: 10.1186/s13640-021-00561-7.

[34] C. Zhang, B. Liang, S. Ma, C. Xiang, Z. Bai, and J. Wang, "Research on the Attribute Extraction Algorithm of People and Vehicles Based on Video Structure," *Journal of Physics: Conference Series,* vol. 1802, no. 3, p. 032081, 2021/03/01 2021, doi: 10.1088/1742-6596/1802/3/032081.

[35] Noor M. Hashem, Heba Kh. Abbas, and H. J. Mohamad, "Detection and Recognition of Car Plates in Parking Lots at Baghdad University," *Iraqi Journal of Science,* vol. 64, no. 2, pp. 1018-1029, 02/28 2023, doi: 10.24996/ijs.2023.64.2.43.

[36] W. Zhang and W. Sun, "Research on small moving target detection algorithm based on complex scene," *Journal of Physics: Conference Series,* vol. 1738, no. 1, p. 012093, 2021/01/01 2021, doi: 10.1088/1742-6596/1738/1/012093.

[37] R. Awad, A. A. Al-Zuky, A. H. Al-Saleh, and H. J. Mohamad, "Enhance Video Film using Retnix method," *Journal of Physics: Conference Series,* vol. 1003, no. 1, p. 012124, 2018/05/01 2018, doi: 10.1088/1742-6596/1003/1/012124.