Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

Computer Vision System For Backflip Motion Recognition in Gymnastics Based On Deep Learning

Ahmed Saadi Abdullah[a], Khalil Ibrahim AlSaif[b]

[a]*Department of Computer Science, College of Education for Pure Science ,University of Mosul/ Mosul,Iraq , Email: ahmed.20csp73@student.unmosul.edu.iq*

[b]*Department of Computer Science, College of Computer Science & Mathematics,University of Mosul/ Mosul ,Iraq,Email: khalil_alsaif@uomosul.edu.iq*

ARTICLE INFO

Article history:
Received: 10/01/2023
Rrevised form: 25/02/2023
Accepted : 28/02/2023
Available online: 31/03/2023

Keywords:

Computer Vision System,

Deep Learning,

Object detection

ABSTRACT

Reliance on computer vision systems in the sports field is one of the very important topics, which are of high importance, especially in the arbitration process or evaluating the accuracy of the player's performance of the movement. It is better to rely on computer vision systems that are more accurate in the arbitration process. In this article, a method was presented to distinguish one of the important movements of the gymnastics player, by relying on deep learning techniques. The dataset was built based on high-quality video clips found on YouTube for tournaments held from the period 2018-2022, due to the absence of The dataset available. This data was divided into three sections: 70% for training, 10% for validation, and 20% for testing. Two models of the convolutional neural network yolov7 and yolov5 were trained, and the results obtained after testing the results of the models show that the seventh version was the best , Recall, Precision and Mean Average Precision criteria were adopted to evaluate the performance of these technologies.

MSC..

https://doi.org/10.29304/jqcm.2023.15.1.1162

1. Introduction

Computer vision is one of the branches of computer science that has begun to show its applications and use in wide areas of life, as it is relied upon in the medical and industrial fields as well as the sports field[1]. Where we now notice clearly the reliance on computer vision systems in the sports field, where these systems are relied upon in the process of locating the player within the matches, as well as the paths of the ball inside the stadium can be determined in matches that use a ball in the game, and there are a set of systems that can be relied upon in Determining the player's performance inside the match By studying the player's movements inside the stadium through these systems, the coach can know the player's performance correctly by analyzing the statistics obtained through these systems. In the team that can perform the plan that the coach thinks of, as well as the coach can analyze the performance of the opposing team and thus determine the weaknesses of the second team, as well as determine the best player in the opposing team, which is better to beware of this player, therefore it must be in the coach's mind To prevent the player from playing [1,2].

∗Corresponding author

Email addresses:

Communicated by 'sub etitor'

After capturing the video of the player inside the field or the ball, the player is then identified inside the field and then distinguishing the player from other players or determining a specific movement of the player or determining the location of the ball inside the field, whereby deep learning techniques are relied on that have the ability to extract the important characteristics of the object To analyze his movements inside the stadium, as deep learning has important features that can identify anyone inside the stadium and thus track his movements or the movement of the ball, as many of these algorithms are used in order to analyze the movement of players in many games, where one of the network models was relied on The bypass nerve in order to identify the player inside the field, after which his movement is tracked, and from there his movement can be studied[3-4].

The gymnastics game is one of the difficult sports in the process of evaluating the player's scores because it needs high accuracy by the arbitrators because the points are given to the player according to the accuracy of the movement implementation and thus relying on computer vision systems and relying on deep learning techniques in the process of distinguishing movement and determining the accuracy of this movement will be more Accuracy in evaluating the scores of this player, the article is divided into seven sections, where the first section provides a general introduction to the importance of computer vision as well as the importance of using these systems in sports reality, while the second section presented a group of articles presented by a group of researchers in analyzing the movements of players and the ball in different players, as for the importance of education The deep and YOLO algorithms has been studied and its architecture is clarified, as it was presented in the third paragraph In the fourth section, the process of tracking objects in the video clips was clarified, while the fifth section presented the proposed method in order to identify and distinguish one of the most important movements of the gymnast, which is the backflip movement. As for the last two sections, the results obtained by applying the proposed method to a group of From the videos of gymnasts, whether this player is an amateur or a professional, as well as the most important conclusions that were reached by analyzing the results obtained.

2.Related Work:

There are many articles that analyze the movements of sports players, and sometimes the tracking of the ball in games that are played with a ball, while some articles presented methods for analyzing the movement of the player's body as well as the trajectories of the ball.

N A Rahmad and others in 2019, presented a method to identify the movement of a badminton player, which is badminton smash, where two models of deep opacity models were relied on, namely Alexnet, GooleNet, where researchers built a special database for this movement, where a group of videos on YouTube were read, and then they were A set of frames were taken for these videos, and thus the database was built, where the database was built, consisting of 528, and this group of images was divided into two parts, one of which is used in the training process, while the other part is used for the testing process, where it was divided by 80% for the training currency and 20% For the test, after applying the proposed method to the video clips in order to determine the accuracy of the proposed method, it was found that the GoogleNet method was the best with an accuracy of 90.3%, while the second method was 86.0% accuracy, and the time required for the training process of the two models shows that GoogleNet took longer than the first process[3].

Nur Anis Jasmin Sufri and others in 2019, presented a method for detecting a badminton player on the court by using one of the Faster R-CNN convolutional neural network models, after collecting a database of badminton player images through videos posted on YouTube for three different tournaments, two of which are badminton tournaments for men's singles and the third It's a men's doubles game. Where the video clip was converted into a set of frames, where 100 images were taken from each video clip, and therefore we will have 300 images, but divided into three databases, where the training of the proposed model was divided into 6 methods, where the model was trained in the first three methods on each One rule, as for the fourth method, the deep learning model used was trained on two bases, one of them is badminton for the single men, and the second rule is for the pairs of men. The reliance on training using three databases was better because all the models in the databases were trained on[5].

Aizreena Azaman and others in 2019 ,presented a method for detecting hit action and non-hit action, and based on four models of the convolutional neural network where (GoogleNet, VGG19, AlexNet, VGG16) were used, these four models were applied to a database consisting of 80 images hardened into 40 hit action images. The other forty of the second movement, this rule was built based on a set of videos on YouTube for one of the badminton tournaments that were held in 2017, where the videos were taken, the length of the clip ranged between 20 to 50 seconds, after which it is transferred to a group of frames. The four models of the convolutional neural network were applied to this rule and it was found that the accuracy that was obtained by relying on GoogleNet reached 87.5%, or the model that reached the best second highest accuracy AlexNet, where the accuracy reached 81.3%, while the remaining two models reached the percentage of accuracy both have 50%[6].

In addition to a group of articles that analyzed another type of sports, the most important of which is basketball, which most of the research depends on the databases obtained through the NBA and a deep learning model is applied to it, where one of the researchers presents a method for analyzing the triple point payment and that By studying the player's movement about throwing the ball in the case of a triple shot[7].other researchers presented a set of methods that relied on computer vision as well as deep learning models from the convolutional neural network models, where it analyzes the player's strategy on the field, which determines the player's strategy as well as determines the best player among the players on the field, in addition to determining the best way to That the team win the matches and the best players can perform the plan. The players on the field want to come up with the best strategy in order to reach the victory[8,9].

3.Deep Learning

Deep learning is a part of machine learning, which falls within artificial intelligence, and Figure (1) shows the relationship between them[10].

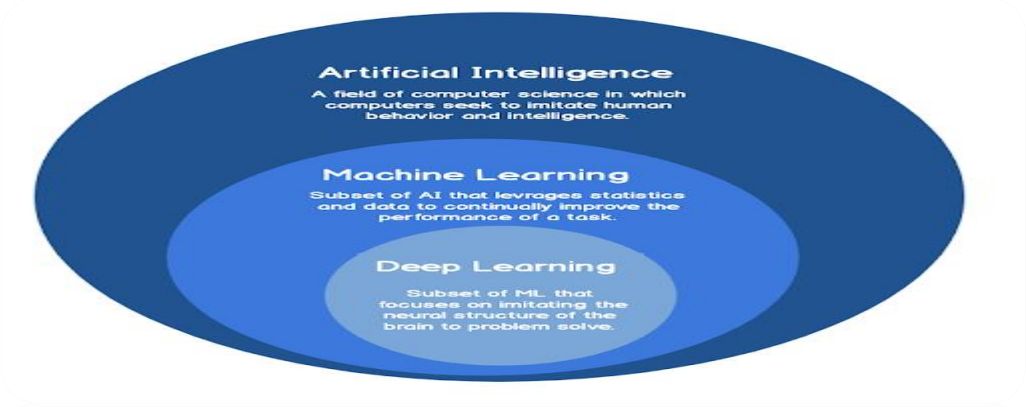


Fig.(1) Artificial Intelligence & Machine Learning & Deep Learning

In general, learning is either supervised or unsupervised. In general, the supervised learning process is used in classification as well as regression, where the inputs have notation, and therefore it can be used to deduce the rules on the basis of which these outputs appear. After training, this machine can identify the things that have been Training on it, but if it is without supervision, then it will not be used in the classification process or in regression, but will be in clustering, where here does not have any information about the outputs, Here only by searching for similarity between these primary data and thus isolating each group of these data with those that are similar to them, whether in shape, color, size, etc[11].

There are many differences between deep learning and machine learning, as table No. (1) shows the most important differences between machine learning and deep learning[11].

Table (1) :Deep Learning VS Machine Learning

Deep Learning	Machine Learning
don't need to extract features manually	**need to extract features manually**
The performance is better on big data	**The performance is good even if the data is not large**
It needs powerful devices because the mathematical operations are complex	**It does not have to be powerful hardware because the computational operations are less complex than deep learning**
The execution time is large because the calculations involved are many, and therefore the time will be large	**Execution time is less than deep learning**

One of the most important deep learning algorithms is YOLO (You Look Only Once), It is one of the most important deep learning algorithms that appeared for the first time in 2015 by Joseph Redmond, after which this algorithm remained in a state of continuous development, as many versions of this algorithm appeared[12], Since this algorithm is detecting the object through one stage, and therefore this algorithm will be very fast and therefore can be exploited in real-time applications, Figure (2) shows the architecture of the fifth and seven version[13].

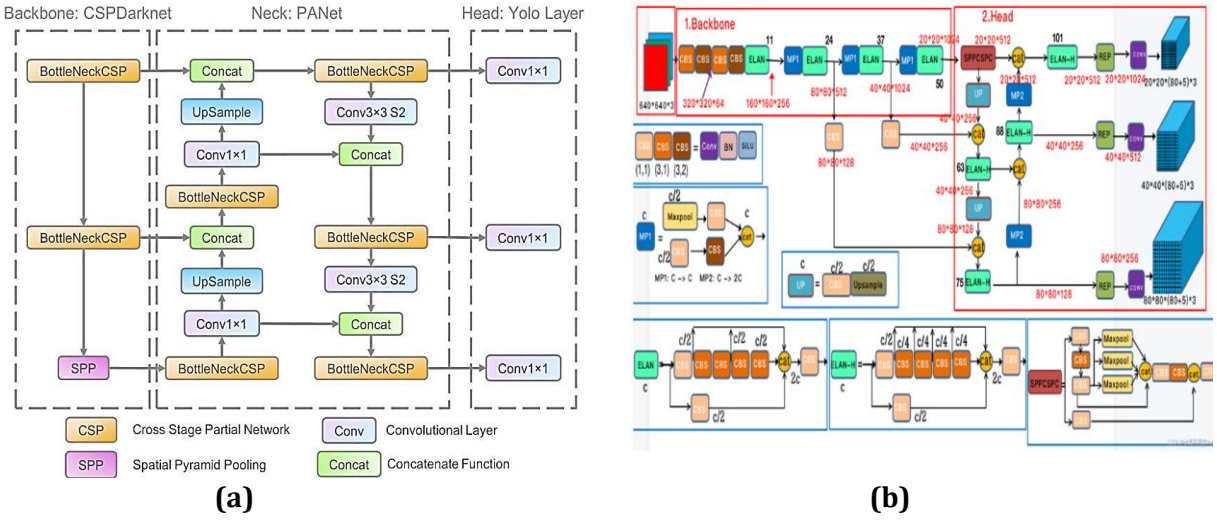


Fig.(2) (a) YOLOv5 architecture , (b) YOLOv7 architecture

4.Proposed System

In this paragraph, we will discuss the steps for implementing the proposed system to detect and distinguish one of the gymnast's movements.

- **first stage,** a dataset is created. Gymnastics is one of the least known sports when compared to other sports such as football or basketball. Therefore, the presence of data for this game will not exist, so the dataset was built by relying on a group of video clips on YouTube. And a group of clips that were taken directly for a group of players from the College of Education, where 400 pictures of this movement were obtained in different directions.

- **second step**, a preliminary image processing group is performed. The contrast of the images is improved by relying on digital image processing techniques, in addition to resizing the image according to the proposed model . divided into three groups, 70% for training, 10% for validation, and the rest for testing.

- third step , Two models of the convolutional neural network were trained, based on the seventh and fifth versions of YOLO, as this model is considered the fastest and most accurate convolutional neural network model in detecting and distinguishing objects.

-fourth step, At this stage, the system is tested by distinguishing and identifying this movement from gymnastics. In the last stage, the performance of the proposed algorithms is measured, as a set of criteria that are used to measure the efficiency of the algorithm are relied upon.

1- Precision:

Variance is computed by subtracting the standard deviation of the sample from the sample mean, as opposed to bias, which is determined by comparing the standard deviation of the collection to the object's known value. What determines whether a classifier's assertion that a collection of records belongs to a positive class is true[14-15].

$$\boldsymbol{precision} = \frac{True\ Posiyive}{True\ Positive + False\ Positive} * 100\% \dots\dots\dots\dots\dots\dots\dots\dots (1)$$

2- Recall:

Recall, as its name suggests, is a measurement of how well a classification system can predict the quantity of positive instances.[14-15]

$$\boldsymbol{Recall} = \frac{True\ Posiyive}{True\ Positive + False\ Nagative} * 100\% \dots\dots\dots\dots\dots\dots\dots\dots (2)$$

3- Mean Average Precision (mAP):

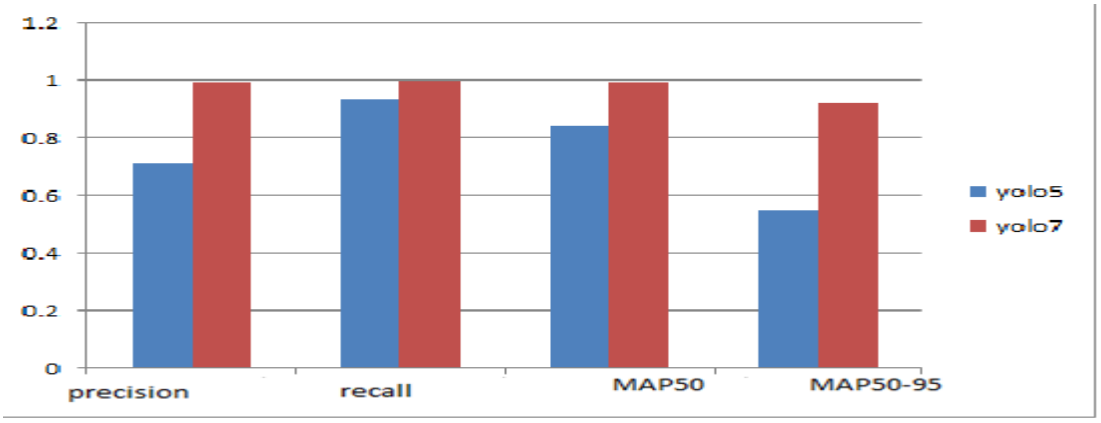
In order to evaluate various object detection models, including YOLO, mean average precision (mAP) is used. A score is given after comparing the detected box to the ground-truth box. The accuracy of the model's forecasts increases with score perfection.However, the accuracy-recall curve can be condensed into a single figure using the average precision (AP) metric.

To calculate your AP, use the following formula. In a cycle that iterates across all precisions and recalls, the difference between the current and following recalls is calculated and then multiplied by the current precision. In other words, the AP is the weighted cumulative sum of precisions that are unique to a threshold, where the weight denotes the relative increase in recall.[15]

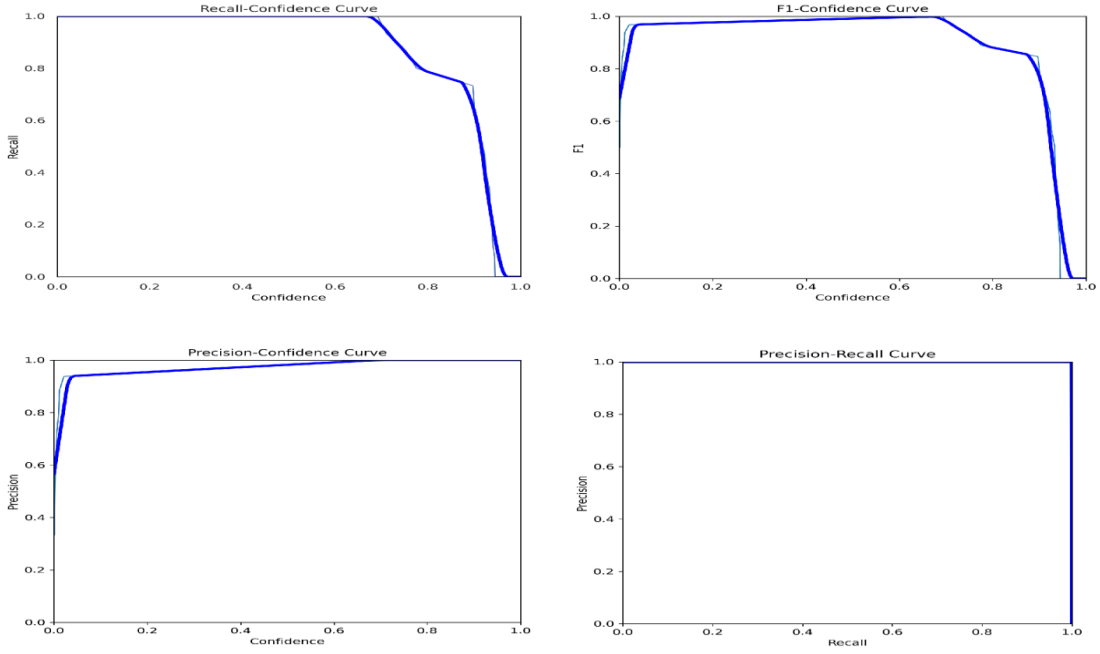
$$\textbf{Mean Average Precision} = \frac{1}{\textbf{number of classes}} \sum_{k=1}^{k=n} The\ Average\ Precision, of\ class\ k. \quad (3)$$

5- Result

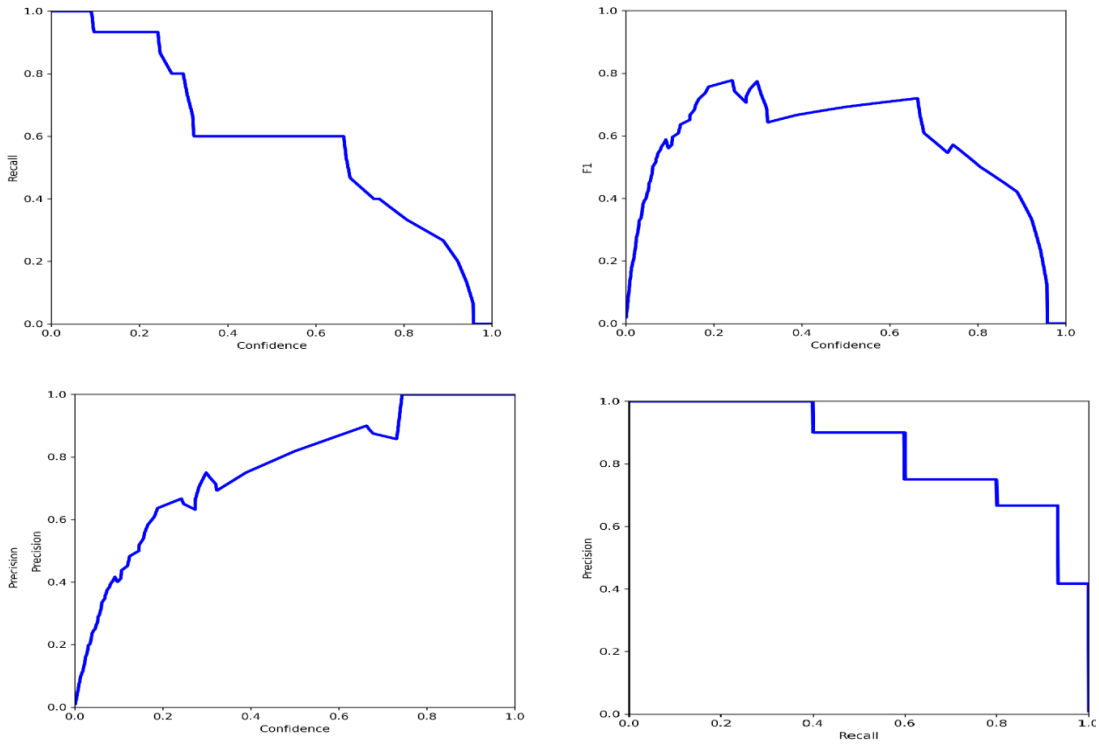
In this paragraph, we will discuss the results obtained from the use of these two models of CNN in order to identify and distinguish one of the most important movements of the gymnast, and the results were as shown in Figures(3-5).



Fig(3): performance YOLO5 VS YOLO7



Fig(4) performance of YOLO V7



fig(5) performance of YOLO V5

It is clear from the figures that represent the results obtained with the process of distinguishing and discovering one of the most important movements of the gymnast, that relying on the yolov7 algorithm gave better results than relying on the fifth version, and the accuracy reached approximately 97%, while relying on the fifth version is approximately the amount of accuracy 84%, but the time it took to train the algorithm for the fifth version was less than the time it takes for the seventh version, and this is due to several reasons. First, the size of the input image for the fifth version is 416 * 416, while the seventh version is 640 * 640, so the processing process will take time In addition, the processes are much longer in the seventh version.

6.conclusion

The process of evaluating the performance of a player in a specific movement is of utmost importance in some sports games in which the outcome of the game depends on the amount of skill that the player has in making the movements, unlike some games in which the result depends on the number of goals such as football, sometimes there are errors Arbitration may be the result of a lack of experience on the part of the referee or the enemy of clarity of vision, and therefore the arbitration process is incorrect. The gymnastic game is considered one of the games that evaluate the player on the skill in the player's performance of the movement. Relying on computer vision systems supported by the deep learning algorithm is better in evaluating the player's performance than relying on human arbitration. Where it is clear from the results that the results that appeared from the seventh version were better, as the accuracy reached 97%, and therefore relying on these algorithms to discover and distinguish the movements of the gymnast player will be appreciated for the player better, also the time taken by the fifth version is less than the seventh version, Because the algorithm of the seventh version is more complex than the fifth version.

References

[1] Graham Thomas, Rikke Gade, Thomas B. Moeslund and Adrian Hilton," Computer vision for sports: current applications and research topics", Computer Vision and Image Understanding, Elsevier,2019.

[2] B. Chakraborty and S. Meher. "A real-time trajectory-based ball detection-and-tracking framework for basketball video" Journal of Optics, , 2013.

[3] N A Rahmad , M A As'ari , K Soeed1 and I Zulkapri," Automated badminton smash recognition using convolutional neural network on the vision based data", Sustainable and Integrated Engineering International Conference 2019 .

[4] Md Nafee Al Islam, Tanzi Bin Hassan, Siamul Karim Khan," A CNN-based approach to classify cricketbowlers basedontheir bowling actions", EEE International Conference onSignal Processing, Information, Communication & Systems,2019.

[5] Nur Azmina Rahmad , Nur Anis Jasmin Sufri , Nurul Hamizah Muzamil , Muhammad Amir As'ari," Badminton player detection using faster region convolutional neural network" Indonesian Journal of Electrical Engineering and Computer Science,2019
.

[6] Aizreena Azaman and others," Recognition of Badminton Action Using Convolutional Neural Network", Indonesian Journal of Electrical Engineering and Informatics,2019.

[7] Peng Yao," Real-Time Analysis of Basketball Sports Data Based on Deep Learning", Complexity,2022.

[8] Rajiv Shah, Rob Romijnders," Applying Deep Learning to Basketball Trajectories" Neural and Evolutionary Computing,2016.

[9] Leili Javadpoura, Jessica Blakesleeb, Mehdi Khazaelib and Pete Schroedera," Optimizing the best play in basketball using deep learning", Journal of Sports Analytic,2022.

[10] Ajeet Ram Pathak , Manjusha Pandey , Siddharth Rautaray ," Application of Deep Learning for Object Detection" Procedia Computer Science ,2018.

[11] Milad Vazan ," Deep learning: principles, concepts and approaches", Miad andishe .2021.

[12] M. Maity, S. Banerjee, and S. S. Chaudhuri, "Faster r-cnn and yolo based vehicle detection: A survey," in 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021.

[13] Muhamad Munawar Yusro ,Rozniza Ali and Muhammad Suzuri Hitam ," Comparison of Faster R-CNN and YOLOv5 for Overlapping Objects Recognition", Baghdad Science Journal,2022.

[14] S. Haghighi, M. Jasemi, S. Hessabi, and A. Zolanvari, "PyCM: Multiclass confusion matrix library in Python," J. Open Source Softw., 2018.

[15] C. S. Hong and T. G. Oh, "TPR-TNR plot for confusion matrix," Commun. Stat. Appl. Methods, , 2021.

[16] S. Visa, B. Ramsay, A. L. Ralescu, and E. Van Der Knaap, "Confusion matrix-based feature selection.," MAICS, 2011.

[17] R. Raj, S. S. Nagaraj, S. Ritesh, T. A. Thushar, and V. M. Aparanji, "Fruit Classification Comparison Based on CNN and YOLO," in IOP Conference Series: Materials Science and Engineering, 2021,