American International Theism University Scientific Research Journal (AITUSRJ)
ISSN (E): 2578-3882**, ISSN (P):** 2578-3874

10.63094/AITUSRJ.25.4.1.5

Integrating Deep Learning for Object Manipulation: A 7-DOF Robotic Arm Perspective on Grasping

Syed Rizwan
Department of Computer Science, IQRA University, Karachi, Pakistan
syed.rizwan@iqra.edu.pk

Benish Fayyaz
Department of Computer Science, IQRA University, Karachi, Pakistan
beenish.fayyaz@yahoo.com

Muhammad Zubair
Department of Computer Science, IQRA University, Karachi, Pakistan
zubair@iqra.edu.pk

Azfar Ghani
Department of Computer Science, IQRA University, Karachi, Pakistan
azfarghani@iqra.edu.pk

Syed Umarullah Hussaini
Department of Robotics and Artificial Intelligence, SZABIST University Karachi, Pakistan
syed.umarullah@szabist.edu.pk

Abstract – The Robotic arm with 7-Degree of Freedom (DOF) is extensively used in numerous industrial applications. However, its precision and control need further improvement for optimum results in various generalized applications. This paper presents a novel approach to improve the manipulation capabilities of a 7-DOF robotic arm by integrating the YOLOv7 object detection model and a Deep Reinforcement Learning (DRL) framework for control. YOLOv7 is employed to provide real-time perception, enabling accurate object recognition, while the DRL algorithm optimizes control by adapting to the dynamic environment of the robotic arm. The DRL algorithm learns through trial and error, adapting to the specific dynamics of the robotic arm and its environment. As a result, improved precision, stability, and adaptability were observed across various tasks. The primary contribution of this work is the optimization and integration of YOLOv7 with a Raspberry Pi, facilitating efficient and real-time object manipulation even on resource-constrained hardware. The proposed algorithm was trained on diverse datasets, enabling the system to generalize effectively across multiple objects and real-world scenarios. Extensive experiments, including repeated trials under varying conditions, demonstrated significant improvements in grasping accuracy and manipulation performance compared to traditional control methods. The system achieved a validated accuracy rate of 94%, supported by statistical analysis and confusion matrix evaluation, confirming its robustness and reliability. These results highlight the potential of intelligent robotic arms to perform complex tasks with high precision and adaptability autonomously.

Index Terms – DOF Robotic Arm, Deep learning, Object Detection, Object Manipulation and grasping, YOLO

I. Introduction

In addition to the era of development, new technology has made robotics and automation, which is one of the other most crucial things that contribute much to the industry's growth. As a result, humanity wants to find intelligent and effective solutions and develop complex systems that enable robots to operate like humans and not like machines. The 7-Degree of Freedom (7-DOF) robotic arm illustrates one of the most revolutionary discoveries.

From the past to this period, different industries have been undergoing hard times compared to now, when 7-DOF robotic arms are fully integrated. These problems concern the hard work and constricted space, regular malfunction, safety concerns, the physical limit of humans, and the risk of human control. The demonstration of these robot arms changed the workability of the ground entirely, and thus, the difficulties were alleviated. In the COVID-19 outbreak, the significance of this technology has been taken to a new level as the capabilities of AI-equipped, 7-Degree of Freedom robotic arms in pick and place tasks have been realized. These machines bring the state by providing uninterrupted services while maintaining the highest concentration on the hygiene program, reducing human interactions, and filling the workforce gaps. COVID-19 demonstrated that production acceleration is key to securing people's health and continuing essential processes.

These robotic arms are also adopted in multiple scenes as they are task-oriented and can operate in hazardous areas. They can replace manual tasks such as long lifting of heavy objects, water duplicates tasks, increase efficiency in multifaceted jobs, ensure safety requirements, and so on. Speaking of the industrial process, they (robots) mechanize the job of painting machines or products to a great extent. Similarly, they are assigned to keep the store shelves packed by retrieving, classifying, and sorting the products taken out from the distribution systems to satisfy customer needs.

A 7-DOF robotic arm has similar functions to a human arm. This mechanism uses AI techniques to execute complex tasks with precision and versatility. It involves object manipulation and pick-and-place operations using an advanced AI. It can independently identify, grasp, and relocate objects skillfully and efficiently. This paper involves the innovative fusion of the 7-DOF robotic arm and AI, exploring its transformative impact on industrial automation, logistics, and other sectors.

As robotic technologies evolve and industries progressively integrate into interconnected networks, the scope of 7-DOF robotic arms continues to broaden. This expansion enables novel applications and innovative operational models.

The subsequent sections of the paper are organized as follows. Section II defines the background study of the system. In section III, different models of object detection are compared. Subsequently, the methodology of the system is described in section IV, in which a block diagram of the system and the object detection model are given. Following this, sections V and VI interpret the testing and results of the system. In section VII, the paper is concluded.

II. Literature Review

Computer vision is a broad field divided into processing, analyzing, and understanding images to turn complex visual data into numerical or symbolic information [1-3]. Therefore, computer vision is an AI working on high-level visual processes like electro-visual perception and comprehension [4]. Many applications utilize computer hardware pre-installed for performing specific tasks, and the more recent ones are based on learning capabilities [5-7]. In this case, changing visual data to different symbols enables the system to make valid conclusions. Image understanding includes reaping invaluable information from images by applying mathematical calculations, physical theory, statistics, and learning models. Detail-detecting segmentation that aims at color and texture is crucial for navigating applications like edge detection and region-based segmentation [8]. Computer vision mechanisms consist of a process in which all the tasks related to visual perception are automated, thus merged into envision operations. This area of research deals with images in general (e.g., videos and medical scans), the extraction from which is often computer-based and thus takes on various theoretical forms. The key objective is the design of computer vision systems that incorporate these principles and models [9]. The trend is to develop intelligent systems consisting of computer eyes with robotic arms. One of the strategies is to learn through the points of view of two or more images that can enable a robot to experiment with 87.8% accuracy [7]. In another study, a computer vision technique was used to transfer control of a robot arm to a colored bottle stopper on each robot joint that was recognized by visual analysis. In two studies, robots imitated rock, paper, and scissors, playing for the opponent using the opponent's hands using the trial pattern approach [10]. In the other experiment, the robot analyzed the opposing hand using computer vision, closed the hand grip using computer vision, and won in the end [11]. In another work, the movements of the human arm originated by themselves, and they could then control the movements of the robot arm in another place at a distance using a vision system [12]. The intelligent autonomous robot systems with computer vision performed autonomously for object recognition based on the shape contour, size, and color [13]. Another project exhibited the artistic robotic arm, which could recognize a randomly placed thing, pick it up, and then move it through a computer vision interface [14].

A range of robotic arms is currently used in AI research, each with distinct features and design considerations. This section will explore some recent and widely recognized robotic arms. The Barrett WAM [15-16] is a cable-driven robot known for its smooth and high-speed operation, boasting a top speed of 3 m/s and repeatability within 2 millimeters. The Meka A2 arm [17] is designed with series-elastic components for human interaction. Other robots with series-elastic arms include Cog, Domo, Obrero, Twenty-One, and the Agile Arm [18-22]. These robots utilize different mechanisms for their series elasticity, with various gearheads such as harmonic drives, planetary gearboxes, and ball screws. While these arms have lower control bandwidth due to series compliance, this hasn't significantly limited their use in manipulation research. Stanford has developed human-safe arms using a macro-mini approach, combining series-elastic mechanisms with small motors to enhance bandwidth [23-24]. The PR2 robot [25-26] features a unique passive gravity compensation system, allowing the arms to move freely in any configuration. The DLR-LWR III arm [27], Schunk lightweight Arm [28], and Robonaut [29] employ motors mounted directly to each joint with harmonic drive gearheads for fast and backlash-free motion. These arms offer slightly higher payloads compared to others, ranging from 3 to 14 kilograms. While not designed for human safety due to their larger flying masses, some demonstrations incorporate force/torque sensors for collision detection. Among commercially available robotic arms, costs are generally high, exceeding $100,000. However, there are affordable examples that can be used in research. The Dynamaid robot's arms [30] utilize lightweight and compact Robotis Dynamixel AI servos. Despite its human-scale workspace, it has a lower payload of 1 kg and a total cost of at least $3500. The KUKA youBot arm [31] is a 5-DOF arm with a small work envelope, repeatability of 0.1 mm, and a payload of 0.5 kg, sold for 14,000 euros.

Robot controllers fall into two categories: combined model-based and model-free mechanisms. Model-based controllers relate robot dynamics, environment model, and more, including the robot's transport dynamics, factors of distracting elements, and gravity. Researchers are working on predictive controllers that can replace the closed-loop and open-loop models in the face of model uncertainties [32-41]. The latest streams of research, which exhibit adaptive controllers with filtering and admittance control, have been proposed [42]. Factor graph kinematic models [38] were constructed based on learning by observation [43]. Decision systems based on fuzzy logic were examined for pick-and-place tasks [44], and systems for human-robot interaction were researched [45]. The model-free controllers attain control without models through the learning processes and without detailed models [47]. Trends show that Machine learning is picking up the space in advanced control systems with model-free and model-based approaches [48-49]. Some model-free controllers apply impedance to the force feedback [50-51] while others base it on an observer data-driven method [52]. We employ reinforcement learning to engage model-free control [52-55]. In summary, model-based controllers have to deal with uncertainty in robotics applications, and model-free ones with successfully saying how to act the robot without a model. By applying both types, the pilot gains skill to improve the accuracy of the movement.

III. Comparison of Object Detection Models

The YOLO (You Only Look Once) family of models has significantly evolved over the years, with each version introducing key innovations to improve performance. Table I shows a comparative summary of the main features and improvements in each version:

TABLE I Comparative Summary of YOLO Models

Model	Key Features	Improvements	Limitations
YOLOv1	Single-stage detection.	Fast and efficient, integrates both localization and classification.	Struggled with small object detection and poor handling of object overlap.
YOLOv2	Introduced anchor boxes, multi-scale detection.	Enhanced object size detection, extended detection capabilities with YOLO9000.	Still faced challenges with small objects and class imbalance.
YOLOv3	Multi-scale, different kernel sizes.	Improved feature extraction, multi-label classification.	Slower than later versions.
YOLOv4	Optimized for GPUs and real-time processing.	Balanced speed and accuracy, practical for real-world applications.	Requires more powerful hardware for optimal performance.
YOLOv5	Simplified architecture with anchor-free design.	Optimized for real-time performance, better accuracy-speed trade-off.	Not as advanced in performance compared to YOLOv7.
YOLOv6	Bi-directional Concatenation (BiC) blocks.	Excellent accuracy and speed for real-time tasks.	Less widespread adoption in research.
YOLOv7	High real-time performance, optimized for edge and cloud.	Achieves top accuracy (56.8% mAP), optimized for mobile GPUs.	Full performance demands higher computational resources.

A. *Customization Of Yolov7 For Robotic Manipulation*

In this study, YOLOv7 was customized to meet the specific requirements of real-time object manipulation with a 7-DOF robotic arm using a Raspberry Pi as the computing platform. The customization focused on several key aspects:

Model Optimization

YOLOv7 was optimized for faster inference and reduced model size to accommodate the hardware constraints. This ensured the model could operate efficiently on the Raspberry Pi while maintaining high detection accuracy for real-time tasks.

Seamless Edge and Cloud Deployment

The adaptability of YOLOv7 to run on both mobile GPUs and cloud platforms was leveraged to enable flexible deployment across different environments, whether local or remote processing was needed.

Real Time Object Detection

With a 56.8% mAP, YOLOv7's high precision enabled the robotic system to accurately identify and manipulate objects in real-time, even in dynamic and cluttered environments.

These customizations made YOLOv7 a powerful tool for enabling efficient, autonomous object manipulation in real-world robotic applications, overcoming the common limitations of computational resources typically found in similar systems.

IV. Methodology

Automation of tasks like picking and placing objects with a 7-DOF robotic arm is always associated with specific processes like hardware design and integration of AI for object detection. The procedure is outlined in the following steps:

A. *Hardware Design And Construction*

The actuation process for the 7-Degree of Freedom (7-DOF) robotic arm starts by assembling a robotic kit after strategically placing the servo motors for accurate movements, then integrating a Raspberry Pi circuit to enhance the computational power. The servo motors are the key to obtaining precise and appropriate motions. Moreover, a robotic tank with DC motors and an ultrasonic sensor is also part of the set to simplify the robotic arm's positioning. The ultrasonic sensor shows real-time data informing a high level of the sensor. A camera is added to meet the demand after training the model object detection on a Raspberry Pi 4 B base. Finally, the robotic kit components and robotic tank kit parts are combined to develop an ultimate integrated system. Utilizing the data from the camera and the ultrasonic sensor, this is one cohesive system that permits the robotic arms to identify objects and orient them to their targeted positions. The hardware architecture of my system is shown in Figure 1.

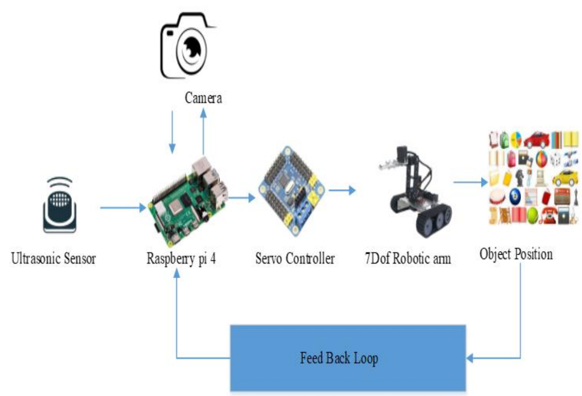


Figure 1 Block diagram of the system

Raspberry Pi 4

The Raspberry Pi 4 is a small single-board computer that acts as the system's brain. It can handle sophisticated algorithms like computer vision. The Broadcom BCM2711 quad-core Cortex-A72 64-bit SoC, clocked at 1.8GHz, and 4GB LPDDR4-3200 SDRAM are the main components of the Raspberry Pi 4. Dual-band WiFi, Bluetooth 5.0, Gigabit Ethernet, and several USB ports are some of its connection capabilities. The gadget can render OpenGL ES 3.1 graphics and generate high-resolution videos using hardware-accelerated video decoding. It works in the 0 to 50 degrees Celsius temperature range and provides expandable storage via micro-SD.

Pi camera module v2

Another well-known product of the Raspberry Pi Foundation, Pi Camera, is perfect for applying image effects, picture taking, and video recording, supplemented by an exclusive visual input function provided by the system. While the LSPC has made efforts to minimize the environmental impact of its products, the Pi Camera Module 2 is a dedicated small-sized image solution for use with the Raspberry Pi. It weighs 3 grams and measures 25 x 24 x 9 mm. The 8MP resolution can accommodate crisp pictures and film at 1080 x 1920 at 47 FPS (frames per second). At the pixel distances of 1.12 µm x 1.12 µm, it ensures high-quality images regardless of the weather conditions. This is called adaptability and calls for a high interest in electronics and hardware programming, monitoring, and spying.

6DOF manipulator alumni robot arm kits base w/servo

TABLE II 6DOF Robotic Arm Kit

Quantity	Items
5 X	Multifunctional bracket
4 X	Long U-shaped bracket
1 X	L-shaped bracket
4 X	Cup bearing
1 X	Alloy mechanical claw
3 X	U-beam bracket
6 X	25T Plastic Servo Arm plate
6 X	MG996 R 180 Degree Metal Gear Servo Motor
1 X	Screw nuts

The mobile robotic arm has a gripper mechanism plus 6 DOF, the degree of robotic movement. Made from a tested, high-strength aluminium frame, it serves for numerous practical applications in both educational and industrial settings due to this fact. The details of all parts are available in Table II.

Servo controllers

A servo controller is a device or system used to control the movement of servo motors. Servo motors are special types capable of precisely controlling position, velocity, and acceleration. They are commonly used in robotics, automation, remote-controlled vehicles, and industrial machinery applications.

A servo controller typically consists of a microcontroller or a dedicated control circuitry that generates control signals to drive the servo motor. These control signals usually consist of pulse-width modulation

(PWM) signals, where the width of the pulse determines the position of the servo motor's shaft.

Vehicle

The robotic arm kit, capable of lifting heavy-duty transport vehicles jointly, has impressive cargo capacity. The vehicle system is based on a tank chassis, which gives it the necessary stability, strength, and mobility on land. The system can increase flexibility and resilience through its tank chassis design, improving its ability to drive in difficult terrain.

DC motor driver

A DC motor driver is an electronic device that creates and controls the speed and direction of a DC motor. The DC motors can be found in diverse fields, from personal recreational projects to professional jobs like machine automation and robotic systems. The DC motor driver gives the motive and all controlling signals to drive the motors efficiently and safely.

Ultrasonic sensor

An ultrasonic sensor is a device that emits ultrasonic sound waves and detects the echo reflected back from objects in its path. These sensors are commonly used for proximity sensing, distance measurement, and object detection in robotics, automation, security systems, and automotive applications.

B. Object Detection Model

Tasks related to object detection in computer vision include designing systems that use cameras for target tracking, surveillance, or smart vehicles.

Difficulties

The Raspberry Pi devices with limited resources are one of the difficulties in deploying object detection models. Among the principal difficulties are:

- Restricted processing capacity

Regarding computing power, the Raspberry Pi marches further behind desktop computers and powerful servers. This limitation may lead to slower inference times as it influences the productivity and the speed at which these complicated object recognition programs churn out outputs.

- Limited memory

In products like the Raspberry Pi, memory constraints sometimes don't allow using big models. The feature parameters and temporary operations associated with object identification, especially in deep learning models, typically require high amounts of memory. It has to be kept in mind that the models must be configured as this memory limitations problem will not occur without losing their functionality.

- Power consumption

Because energy efficiency is one of the primary goals, resource-limited devices are often used in such settings. With many computationally complicated object identification techniques, a quick drain on the battery is in order, or if grid interconnection is done, expenses go high.

Methodology

This methodology adapts the YOLOv7 algorithm for custom object detection on the Raspberry Pi.

- Dataset preparation

The Boxes vs. Car Image Classification Dataset was carefully designed to meet the application's specific needs, focusing on two primary object classes: boxes and cars. A total of 5,000 training images were collected, equally divided between the two categories, with each image resized to 224×224 pixels. To improve the diversity and quality of the dataset, multiple augmentation techniques were applied, including horizontal flipping, slight rotations (up to ±10 degrees), minor zooming (±10%), and controlled brightness adjustments (±10%). These augmentations expanded the dataset to over 20,000 images, introducing a variety of angles, lighting conditions, and backgrounds.

The validation and testing datasets each contain 1,500 images, maintaining a balanced class distribution. Before training, all images were pre-processed to ensure that the objects were consistently centered and correctly scaled, minimizing positional variations. The dataset was further strengthened with 500 additional images captured from real-world environments using cameras, complemented by images sourced from online repositories such as Google Images. This synthetic and real-world data combination was intended to improve the model's ability to generalize and maintain high classification performance when applied to diverse, real-life situations.

- Model Configuration

The dataset's specific nuances are used to formulate an architecture that is a custom version of YOLOv7. This model has a Darknet clawback with a bottom of 53 levels and a PANet (Path Aggregation Network) neck module, using the YOLOv3 style head to predict. The dim-to-dots boxes are defined for different scales, and the model is trained with a batch size of 16 using the Adam optimizer with a learning rate of 0.002 for 50 epochs.

- Training

The model's training uses the Darknet framework, and the dataset is separated into three sections, which are commonly referred to as the training, testing, and validation sets. Correct learning and generalization are achieved by thoroughly checking the progress made during training and by the model validation performance.

- Boosting the performance of Raspberry Pi

Implementing various methods designed to leverage the device's hardware capabilities is essential. Given the limited computing power of the Raspberry Pi, optimizing the model becomes critical to strike a balance between efficacy and accuracy. Methods such as quantization, pruning, and knowledge transfer are utilized to reduce the number of parameters and computational demands while maintaining classification accuracy. Quantization involves optimizing the precision of model parameters by reducing the number of bits used to represent them, thereby conserving memory and computational resources. Pruning entails trimming unnecessary connections or parameters from the model, thereby reducing its complexity while preserving desired performance levels. Knowledge distillation involves training a smaller model to mimic the behavior of a larger model. In this process, knowledge is transferred from the larger model to the smaller one, enhancing computational efficiency without sacrificing accuracy.

C. Path Planning and Control

The optimal route planning algorithms determine the perfect path for the robotic arm to balance on and then reach the exact point where the robot will grasp the object. This instantaneous behavior control is done by real-time algorithms, which ensure that motions happen better and the arm follows the rules of dynamics and mechanics.

V. TESTING

The working results of our 7-DOF Robotic arm are as follows:

A. Testing The Arm's Maneuverability

The robotic arm is part of the system, whose main construction is composed of 6 servo motors and 1 DC motor; this makes a 7-DOF (Degrees of Freedom) system. The motors are placed in different parts of the arm's structure, possibly controlling motion like yaw, pitch, roll, elbow bending, and tanking. Every servo can perform joint movements within specific angle ranges, which are defined as the life zone, while the areas with no control signals are known as the dead zone. We conducted experiments to determine these zones with plots shown in Figure 2 and Table III. These plots give a detailed outline of the permitted joint movements and the movement restrictions. In parallel, the DC motor controls the robotic arm to move from one place to another towards the desired point.

TABLE III THE LIFE AND DEAD ZONE OF SERVO AND DC MOTORS

Motors	Life zone	Dead zone
Servo 1	0^0 to 310^0	311^0 to 360^0
Servo 2	40^0 to 320^0	0^0 to 39^0, 321^0 to 360^0
Servo 3	50^0 to 290^0	0^0 to 49^0, 291^0 to 360^0
Servo 4	0^0 to 230^0	231^0 to 360^0
Servo 5	0^0 to 300^0	301^0 to 360^0
Servo 6	0^0 to 180^0	181^0 to 360^0
DC 1	0^0 to 360^0	18^0 to 60^0

B. Testing the System's Functionality

YOLOv7 was deployed on a Raspberry Pi, using their images to perform real-time training for personalized object detection. Systematically, photos of each item from various perspectives were captured, and each item was prepared and documented individually to enhance the model's recognition capabilities.

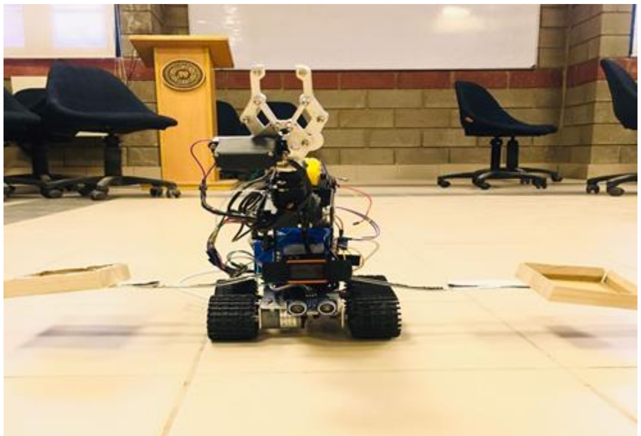
In the tank, which the Raspberry Pi integrates to implement an object detection system, using YOLOv7, and motorizes it with the DC motors for motion. This sends a signal in a straight line, while the probe uses an ultrasonic sensor to detect any obstacles. The tank switches on its camera once it has collected the information from the sensor to detect an object. Furthermore, it correctly decides where the object should be put in the series, either on the left or right sides.

The ultrasonic sensors, reliable in their functionality, do the spatial checking between a vehicle and the identified obstacles in the environment. Moreover, the data includes ground clearance, which the car uses to flatten the ground so it can go through the ground safely. The car is driven until it reaches the target item, such as another vehicle, picks it up, and takes it to the correct destination. For instance, the robotic arm detects an object like a box, lifts it, and repositions it into the box on the left.

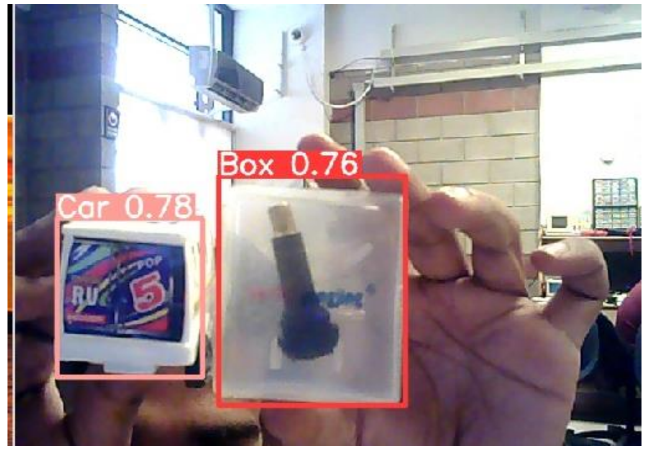
Suppose the robot senses the presence of both objects. In that case, the 7-DOF robot arm first picks up the item with the highest precision by considering prioritization, then loads the other goods shown in Figure 2.

The position and orientation of the object are also detected by the servo motors inside the robotic arm kit, which then uses that information to grasp it properly and place it in the correct position. With an actuator set of 6 servo motors, the 7-DOF robotic arm is ready to perform the task adroitly and accurately. Hence, the robotic arm can adjust its posture and alignment to grasp any object picked by the vision system successfully. As the robot under the control, the robotic arm assists the doctor in doing the job of a great surgeon. The hand of the robotic arm moves the object firmly to the intended location, and skillfully sets down the item at the designated zone without any harm.

American International Theism University Scientific Research Journal (AITUSRJ)
ISSN (E): 2578-3882, **ISSN (P):** 2578-3874

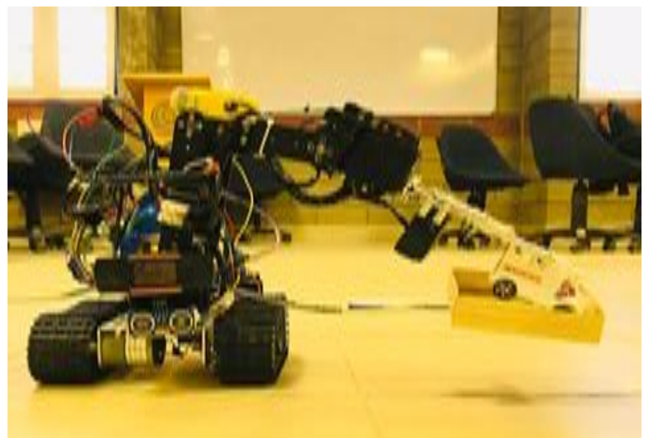


(a) Robotic arm while searching



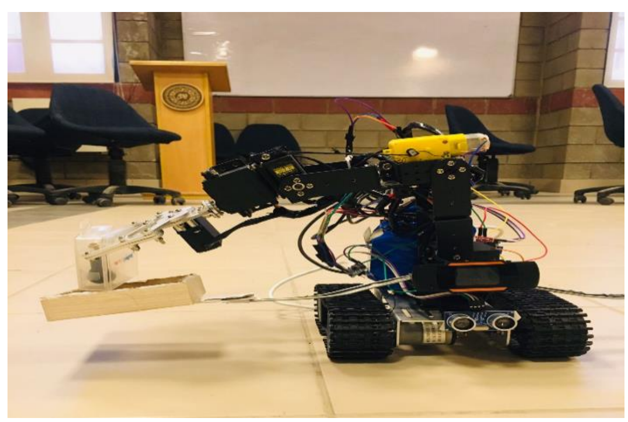
(b) Object detection of both objects

(c) Custom object detection of a car



(d) Grasping the car object

(e) Custom object detection of a box



(f) Grasping the box object

Figure 2 Complete working of a 7-DOF Robotic Arm

C. Testing the Arm's Accuracy

TABLE IV ARM'S ACCURACY

Parameters	7-DOF Robotic Arm
Number of axes	7 axes
Payload	5kg
Weight	18kg
Speed	1m/s
Accuracy	0.94

Table IV represents the arm's accuracy, which summarizes the performance. Each row of the table defines a parameter and its value.

VI. RESULTS

Here are the outcomes derived from the utilization of the 7-DOF robotic arm:

The top is the Training Validation Statistics window, and the mosaics, labels, predictions, and Augmented Mosaics are displayed in the other six windows, respectively, from the first to the sixth. Besides that, the measurement tables and the corresponding charts, such as the PR curve and the confusion matrix, are also there. It depicts the most significant problem as shown in Figure 3.

To ensure that the object detection model is valid and robust, we exhaustively applied all metrics available to verify its precision and stability. These parameters were used to analyze factors such as the bounding box regression (box_loss), the reliability of object prediction (obj_loss), and classification accuracy (cls_loss). It was not only accuracy but also precision that the modeler concerned was the ratio of well-predicted bounding boxes to all predicted ones, since this value gives an idea of how exactly it pinpointed relevant objects. Simultaneously being able to see if the model has detected a substantial part of the bounding boxes correctly, recall exhibits the model's sensitivity in object location within an image.

Moreover, a part of our appraisal is computing a mAP (Mean Average Precision), which considers Intersection over Union (IoU) and is referred to as mAP@0.5. Now we have a perfect overall gauge of the model in identifying objects. To offer a more comprehensive view of model performance, we calculated 'mAP_0.5:'0.95,' the mean of the mAP at various IoUs from 0.5 to 0.95, which is a standard measure for IoUs evaluation. Efficiently, that was information analysis of the metrics related to the object detection system, in general, made it possible to identify the strengths and weaknesses of that system, which, in turn, allowed a detailed and advanced evaluation report of the system.

With its evaluation results, the object detection model has done a great job. The performances were analyzed using multiple metrics, as shown in Figure 4. The F1 confidence curve shows the models precisely, and recall rates indicate their performance in classifying and locating objects with confidence levels ranging from 0.99 to 0.845. With a high confidence level of accuracy maintained while the false positive detection rate is optimized, this is further supported by the precision confidence curve that runs from 1.00 to 0.941. The recall-precision curve is the feature that stands out in the model, and it yields a fantastic map of 0.993 for a confidence threshold of 0.5. This shows the LWIM model can be used to attain accurate and timely diagnoses in multiple cases. In addition, the recall confidence curve reveals the model's flexibility, which suggests that it may perform better or worse at different confidence levels. These data points, which accurately reflect the model's results, verify the model's effectiveness and, thus, make it a tool of choice for applications that require precise and reliable object detection over a range of confidence thresholds.

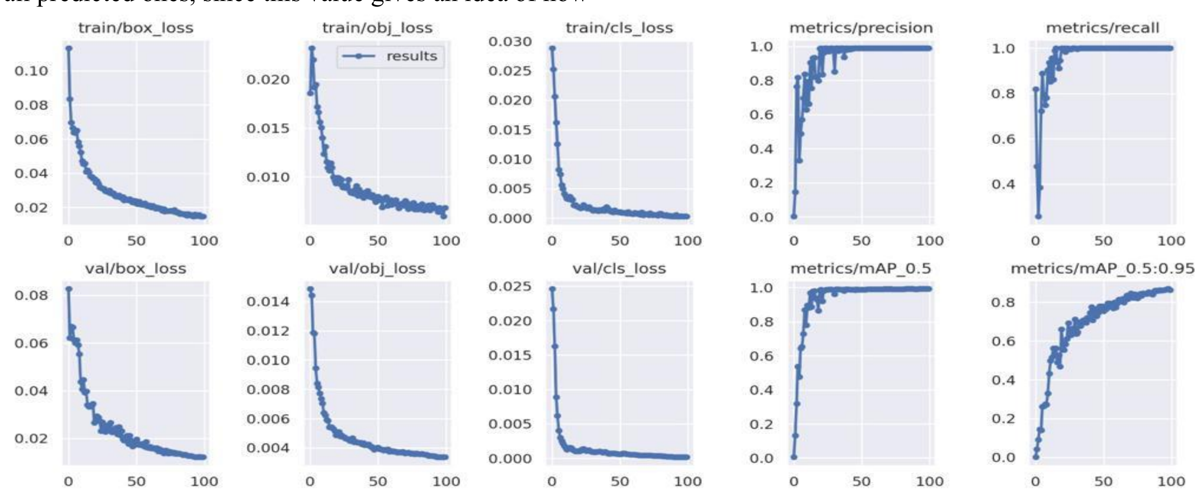
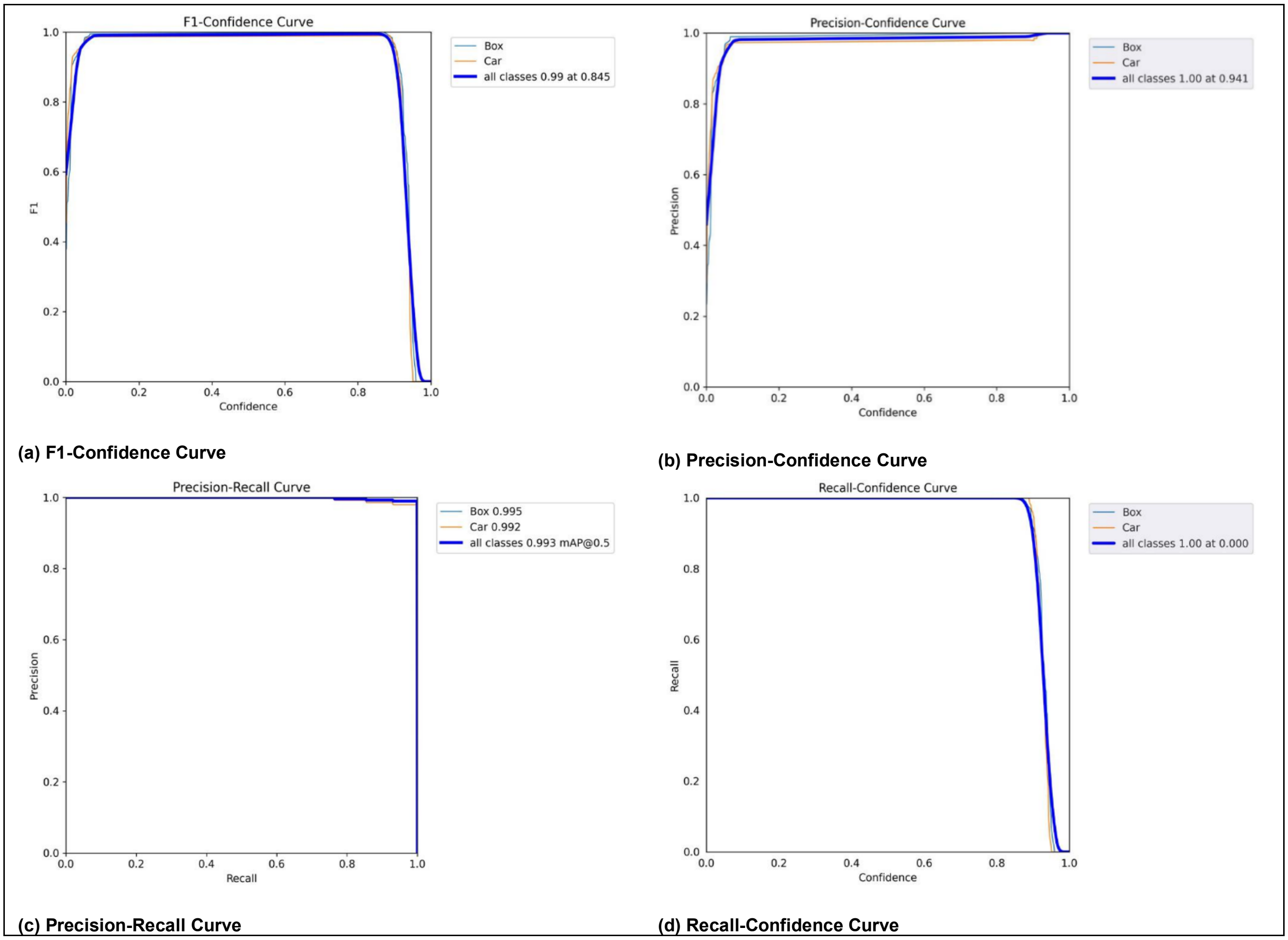


Figure 3 Validation Performance during the training process



(a) F1-Confidence Curve **(b) Precision-Confidence Curve**

(c) Precision-Recall Curve **(d) Recall-Confidence Curve**

Figure 4 Validation Performance during the testing process

VII. CONCLUSION

This study highlights the pivotal role of deep learning in advancing the capabilities of robotic arms. The findings demonstrate that incorporating deep learning techniques significantly improves single-armed robotic systems' grasping and manipulation performance. These results suggest promising directions for future developments in intelligent robotic applications.

Furthermore, the test results of achieved performance characteristics showed the performance of the approach, which is 94% efficient in object recognition. In addition, the system shows a fair precision-recall ratio, especially 99% for each of the two. The system's uniqueness is that it can reach the perception with 100% confidence, implying that it can be relied on when there is a need to decide between grasping and manipulation. The perfect performance parameters indicate a great deal of progress in robot manipulation. The proposed deep learning-based control technique provides a solid core that can be implemented for industrial use.

The adaptability of the robotic arm is proven by the fact that it can handle all objects regardless of their shapes, sizes, and orientations because of its abilities. The system's autonomy, in using deep reinforcement learning effectively, makes it adaptable and creates the possibility of dynamic motion optimization coupled with the assurance of accurate results in complex contexts. In this way, the successful outcome of our study draws a plan for new perspectives of the expert community, and it offers new opportunities for in-depth consideration and full-scale implementation. The integration of high-performing perception and intelligent control illustrated in the proposed development lays the groundwork for developing more robust robotic solutions. These systems will be equipped with universal knowledge and a tremendous ability to preserve the central processes in such a sophisticated environment, helping to develop various industries.

The academic relevance of our work and its practical significance in the context of deep learning in robotics cannot be overstated. This is especially useful for industries that need advanced and intelligent robotic systems for precise manipulation tasks. It serves as a mere reflection of a reality where the smart machines are humbly part of the daily activities, and boldly take successful functions to the next level of intelligence without losing autonomy.

REFERENCES

[1] D. Kulkarni, Computer vision and fuzzy-neural systems. Prentice Hall PTR, 2001.

[2] M. Ebner, "A parallel algorithm for color constancy," Journal of Parallel and Distributed Computing, vol. 64, no. 1, pp. 79–88, 2004.

[3] D. Forsyth and J. Ponce, "Prentice hall professional technical reference," Computer vision: a modern approach, 2002.

[4] C.-Y. Tsai, C.-C. Wong, C.-J. Yu, C.-C. Liu, and T.-Y. Liu, "A hybrid switched reactive-based visual servo control of 5-dof robot manipulators for pick-and-place tasks," IEEE Systems Journal, vol. 9, no. 1, pp. 119–130, 2015.

[5] G. Bradski, A. Kaehler, and V. Pisarevsky, "Learning- based computer vision with intel's open source computer vision library." Intel Technology Journal, vol. 9, no. 2, 2005.

[6] C. H. Lampert, H. Nickisch, and S. Harmeling, "Learning to detect unseen object classes by between- class attribute transfer," in Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009, pp. 951–958.

[7] Saxena, J. Driemeyer, and A. Y. Ng, "Robotic grasping of novel objects using vision," The International Journal of Robotics Research, vol. 27, no. 2, pp. 157–173, 2008.

[8] M. Kazemi, K. K. Gupta, and M. Mehrandezh, "Randomized kinodynamic planning for robust visual servoing," IEEE Transactions on Robotics, vol. 29, no. 5, pp. 1197–1211, 2013.

[9] D. Song, C. H. Ek, K. Huebner, and D. Kragic, "Task- based robot grasp planning using probabilistic inference," IEEE transactions on robotics, vol. 31, no. 3, pp. 546–561, 2015.

[10] R. Szab´o and A. S. Gontean, "Full 3d robotic arm control with stereo cameras made in labview." in FedCSIS Position Papers, 2013, pp. 37–42.

[11] Y. Hasuda, S. Ishibashi, H. Kozuka, H. Okano, and J. Ishikawa, "A robot designed to play the game" rock, paper, scissors"," in Industrial Electronics, 2007. ISIE 2007. IEEE International Symposium on. IEEE, 2007, pp. 2065–2070.

[12] Shaikh, G. Khaladkar, R. Jage, and T. P. J. Taili, "Robotic arm movements wirelessly synchronized with human arm movements using real time image processing," in India Educators' Conference (TIIEC), 2013 Texas Instruments. IEEE, 2013, pp. 277–284.

[13] S. Manzoor, R. U. Islam, A. Khalid, A. Samad, and J. Iqbal, "An open-source multi-dof articulated robotic educational platform for autonomous object manipulation," Robotics and Computer-Integrated Manufacturing, vol. 30, no. 3, pp. 351–362, 2014.

[14] T. P. Cabre, M. T. Cairol, D. F. Calafell, M. T. Ribes, and J. P. Roca, "Project-based learning example: controlling an educational robotic arm with computer vision," IEEE Revista Iberoamericana de Tecnologias del Aprendizaje, vol. 8, no. 3, pp. 135–142, 2013.

[15] B. Rooks, "The harmonious robot," Industrial Robot: An International Journal, vol. 33, no. 2, pp. 125–130, 2006.

[16] Barrett Technology, Inc., "WAM Arm," 2010. [Online]. Available: http://www.barrett.com/robot/products-arm-specifications.html.

[17] Meka Robotics, "A2 compliant arm," 2009. [Online]. Available: http://www.mekabot.com/arm.html.

[18] R. Brooks, C. Breazeal, M. Marjanovi´c, B. Scassellati, and M. Williamson, "The Cog project: Building a humanoid robot," Computation for metaphors, analogy, and agents, pp. 52–87, 1999.

[19] Edsinger-Gonzales and J. Weber, "Domo: A force sensing humanoid robot for manipulation research," in 2004 4th IEEE/RAS International Conference on Humanoid Robots, 2004, pp. 273–291.

[20] E. Torres-Jara, "Obrero: A platform for sensitive manipulation," in 2005 5th IEEE-RAS International Conference on Humanoid Robots, 2005, pp. 327–332.

[21] H. Iwata, S. Kobashi, T. Aono, and S. Sugano, "Design of anthropomorphic 4-dof tactile interaction manipulator with passive joints," Intelligent Robots and Systems, 2005 (IROS 2005), pp. 1785 – 1790, Aug. 2005.

[22] J. Pratt, B. Krupp, and C. Morse, "Series elastic actuators for high fidelity force control," Industrial Robot: An International Journal, vol. 29, no. 3, pp. 234–241, 2002.

[23] M. Zinn, B. Roth, O. Khatib, and J. Salisbury, "A new actuation approach for human friendly robot design," The international journal of robotics research, vol. 23, no. 4-5, p. 379, 2004.

[24] D. Shin, I. Sardellitti, and O. Khatib, "A hybrid actuation approach for human-friendly robot design," in IEEE Int. Conf. on Robotics and Automation (ICRA 2008), Pasadena, USA, 2008, pp. 1741–1746.

[25] K. Wyrobek, E. Berger, H. der Loos, and J. Salisbury, "Towards a personal robotics development platform: Rationale and design of an intrinsically safe personal robot," in Proc. IEEE Int. Conf. on Robotics and Automation, 2008, pp. 2165–2170.

[26] Willow Garage, "PR2," 2010. [Online]. Available: http://www.willowgarage.com/pages/pr2/specs.

[27] G. Hirzinger, N. Sporer, A. Albu-Schaffer, M. Hahnle, R. Krenn, A. Pascucci, and M. Schedl, "DLR's torque-controlled light weight robot III- Are we reaching the technological limits now?" in Proceedings- IEEE International Conference on Robotics and Automation, vol. 2, 2002, pp. 1710–1716.

[28] Schunk, "7-DOF LWA Manipulator," 2010. [Online]. Available: http://www.schunk-modular-robotics.com/left-navigation/service-robotics/components/manipulators.html.

[29] R. Ambrose, H. Aldridge, R. Askew, R. Burridge, W. Bluethmann, M. Diftler, C. Lovchik, D. Magruder, and F. Rehnmark, "Robonaut: NASA's space humanoid," IEEE Intelligent Systems and Their Applications, vol. 15, no. 4, pp. 57–63, 2000.

[30] J. Stuckler, M. Schreiber, and S. Behnke, "Dynamaid, an anthropomorphic robot for research on domestic service applications," in Proc. of the 4th European Conference on Mobile Robots (ECMR), 2009.

[31] KUKA, "youbot arm," 2010. [Online]. Available: http://www.kuka-youbot.com.

[32] Zhang Q, Zhou J, Wang H, Chai T. 2016. Output feedback stabilization for a class of multi-variable bilinear stochastic systems with stochastic coupling attenuation. IEEE Transactions on Automatic Control 62(6):2936-2942.

[33] Liu P, Yu H, Cang S. 2014. Modelling and control of an elastically joint-actuated cart-pole underactuated system.

[34] Liu P, Yu H, Cang S. 2016. Modelling and dynamic analysis of underactuated capsule systems with friction-induced hysteresis.

[35] Liu P, Yu H, Cang S. 2018a. Geometric analysis-based trajectory planning and control for underactuated capsule systems with viscoelastic property. Transactions of the Institute of Measurement and Control 40(7):2416-2427.

[36] Liu P, Yu H, Cang S. 2018b. On the dynamics of a vibro-driven capsule system. Archive of Applied Mechanics 88(12):2199-2219.

[37] Liu P, Yu H, Cang S. 2018c. Optimized adaptive tracking control for an underactuated vibro-driven capsule system. Nonlinear Dynamics 94(3):1803-1817.

[38] Liu P, Yu H, Cang S. 2018d. Trajectory synthesis and optimization of an underactuated microrobotic system with dynamic constraints and couplings. International Journal of Control, Automation and Systems 16(5):2373-2383.

[39] Liu P, Yu H, Cang S. 2019. Adaptive neural network tracking control for underactuated systems with matched and mismatched disturbances. Nonlinear Dynamics 98(2):1447-1464.

[40] Liu P, Neumann G, Fu Q, Pearson S, Yu H. 2018b. Energy-efficient design and control of a vibro-driven robot.

[41] Huda MN, Liu P, Saha C, Yu H. 2020. Modelling and motion analysis of a pill-sized hybrid capsule robot. Journal of Intelligent & Robotic Systems 100(3–4):753-764.

[42] Cao P, Gan Y, Duan J, Dai X. 2019. Passivity-based stable human-robot cooperation with variable admittance control.

[43] Sugiarto I, Conradt J. 2017. A model-based approach to robot kinematics and control using discrete factor graphs with belief propagation. Robotics and Autonomous Systems 91(9):234-246

[44] Toan NV, Khoi PB. 2019. Fuzzy-based-admittance controller for safe natural human–robot interaction. Advanced Robotics 33(15–16):815-823.

[45] Kang G, Oh HS, Seo JK, Kim U, Choi HR. 2019. Variable admittance control of robot manipulators based on human intention. IEEE/ASME Transactions on Mechatronics 24(3):1023-1032.

[46] Abraham I, Handa A, Ratliff N, Lowrey K, Murphey TD, Fox D. 2020. Model-based generalization under parameter uncertainty using path integral control. IEEE Robotics and Automation Letters 5(2):2864-2871.

[47] Popović M, Kootstra G, Jørgensen JA, Kragic D, Krüger N. 2011. Grasping unknown objects using an early cognitive vision system for general scene understanding.

[48] Williams G, Wagener N, Goldfain B, Drews P, Rehg JM, Boots B, Theodorou EA. 2017. Information theoretic MPC for model-based reinforcement learning.

[49] Wu M, Taetz B, Saraiva ED, Bleser G, Liu S. 2019b. On-line motion prediction and adaptive control in human-robot handover tasks.

[50] Lee MA, Zhu Y, Zachares P, Tan M, Srinivasan K, Savarese S, Fei-Fei L, Garg A, Bohg J. 2020. Making sense of vision and touch: learning multimodal representations for contact-rich tasks. IEEE Transactions on Robotics 36:582-596.

[51] Wang H, Wang Z, Wang H. 2019. Impedance control strategy and experimental analysis of collaborative robots based on torque feedback.

[52] Esmaeili B, Salim M, Baradarannia M, Farzamnia A. 2019. Data-driven observer-based model-free adaptive discrete-time terminal sliding mode control of rigid robot manipulators.

[53] Perrusquía A, Yu W, Soria A. 2019. Optimal contact force of robots in unknown environments using reinforcement learning and model-free controllers.

[54] Wu W, Li D, Meng W, Zuo J, Liu Q, Ai Q. 2019a. Iterative feedback tuning-based model-free adaptive iterative learning control of pneumatic artificial muscle.

[55] Zhou Y, Zhang Q, Wang H, Zhou P, Chai T. 2017. EKF-based enhanced performance controller design for nonlinear stochastic systems. IEEE Transactions on Automatic Control 63(4):1155-1162.

[56] Zhang Q-C, Hu L, Gow J. 2020. Output feedback stabilization for MIMO semi-linear stochastic systems with transient optimisation. International Journal of Automation and Computing 17(1):83-95.

[57] Joseph, R & Ali, F. (2018). YOLOv3: An Incremental Improvement. 1804.02767. https://docs.ultralytics.com/models/yolov3/

[58] Bochkovskiy, A, Wang, C & Mark Liao, H. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. 2004.10934. https://docs.ultralytics.com/models/yolov4/

[59] Jocher, G. (2020). Ultralytics YOLOv5. https://doi.org/10.5281/zenodo.3908559/

[60] Li, C & ET all. (2023). YOLOv6 v3.0: A Full-Scale Reloading. 2301.05586. https://docs.ultralytics.com/models/yolov6/

[61] Wang, Chien-Yao and Bochkovskiy, Alexey and Liao, Hong-Yuan Mark. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. 2207.02696. https://docs.ultralytics.com/models/yolov7/

[62] Jocher, G, Chaurasia, A & Qiu, J. (2023). Ultralytics YOLOv8. https://github.com/ultralytics/ultralytics