International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Precision Robotics Arm System based on Computer Vision

Shripad Bhatlawande[1], Swati Shilaskar[2], Ansari Maaz[3*]

Submitted: 27/01/2024 **Revised**: 05/03/2024 **Accepted**: 13/03/2024

Abstract: This paper presents an integrated approach to real-time object detection and precise control of a robotic arm (RA), addressing the challenge of seamless environmental interaction. The system utilizes the You Only Look Once version 4 (YOLOv4) algorithm for swift and accurate object identification, along with forward kinematics for RA tracking, ensuring accuracy and responsiveness in real-world applications. The innovation lies in combining different convolutional neural network (CNN) architectures while maintaining precision in implementing the control mechanism with an Arduino Uno microcontroller. Initial implementations for amputees are explored, promising enhanced interaction and autonomy. Validation accuracies of 91.78% and 89.92% highlight the system's effectiveness. Ongoing evaluation and dataset diversification are essential for advancement.

Keywords: *Industrial Automation, object detection, object classification, robotic arm, degrees of freedom*

I. Introduction

The integration of robots in various fields is increasingly replacing human labor in repetitive or hazardous tasks. The use of mobile robot bases, including drones, has expanded the capabilities of cooperative manipulation, extending the reach of robots in different workspaces[1]The dynamic evolution of technology has spurred a worldwide emphasis on industrial automation, with robotic systems, particularly robotic arms, serving as pivotal contributors. The convergence of enhanced software capabilities, advanced hardware components, and sophisticated motors has granted robotic arms newfound versatility, enabling them to execute intricate tasks in diverse domains. This transformation spans from the precision demands of industrial applications to the nuanced requirements of activities like badminton, showcasing the adaptability of these robotic systems.[2] Camera pose estimation determines the camera's location and perspective on an object or scene, crucial in robotics, augmented reality, and medical procedures. The study is dedicated to enhancing the accuracy of this estimation, specifically for tasks such as robotic grasping, with methods like homography decomposition and rigid pose estimation being employed. Additionally, this work extends its focus to integrate these advancements into the realm of robotic arm operations, further emphasizing practical applications in the field. [3]Within the medical landscape, the transformative integration of robots into surgical procedures, encompassing fields like robotic arm applications, has ushered in an era of less invasive interventions. Meticulous control over soft tissue dynamics during suturing holds paramount significance for the success of these procedures. This introduces a dual-arm robotic strategy, strategically minimizing dependence on detailed mechanical information, with the explicit goal of automating suturing processes. The emphasis lies in achieving precise and reliable needle insertion, thereby advancing the frontier of surgical automation [4]A robotic arm is employed to control the position and orientation of a transmitting antenna inside the RC, providing a greater number of samples compared to manual methods. The results are compared with conventional mechanical stirrers in terms of different figures of merit. The integration of a robotic arm in the Radio Chamber demonstrates not only superior sample generation capabilities but also emphasizes the potential for advancing experimental outcomes in comparison to traditional mechanical methods. [5]

[1] *Vishwakarma Institute of Technology, Pune, India*
ORCID ID : 0000-0001-8405-9824
[2] *Vishwakarma Institute of Technology, Pune, India*
ORCID ID : 0000-0002-1450-2939
[3] *Vishwakarma Institute of Technology, Pune, India*
ORCID ID : 0009-0004-6516-5049
** Corresponding Author Email: maaz.ansari211@vit.edu*

II. Literature Review

The investigation into cooperative arms, mobile manipulators, and aerial manipulators for collaborative tasks involves utilizing drones as a communication layer for dual-arm robots. Additionally, the proposed control method involves using brain signals, aiming to enhance simplicity, facilitate AI development, and improve accessibility for disabled operators.[6] A smart control strategy for a robotic arm handling water-filled bottles was developed, employing the AIWCPSO algorithm and velocity curve design. The objective was to determine the optimal approach for moving the bottles without spilling. The results highlighted the success of the intelligent control strategy, suggesting its potential effectiveness for similar tasks in the future.[7] A new method for determining camera position relative to a flat target using multiple images and robot position information has been proposed. Their specialized optimizer improves accuracy compared to existing approaches.

Validation with computer-generated and real-world data shows superior results to other methods, including Zhang's stereo camera calibration and the Introspective Multiview Approach [8] Fangxun Zhong et al. present a dual-arm control strategy for accurate robotic needle insertion in minimally-invasive procedures. The approach actively manipulates both the needle and tissue, reducing errors and improving insertion accuracy while addressing target deviation caused by tissue deformation. The study highlights the potential for automated robotic suturing with competitive accuracy compared to manual human execution. [9] In the pursuit of efficient sample generation, researchers devised a 3-D printed robotic arm to construct a source-stirred Radio Chamber (RC).Source stirring outperformed mechanical stirring, offering a smaller stirring volume. However, at low frequencies, antenna movement caused periodic high correlations, and at high frequencies, cable movement introduced uncertainties in antenna efficiency measurements. [10] In the realm of dual-arm exoskeleton robotics, a coordination control method is proposed, emphasizing human impedance transfer skills. The left arm extracts stiffness and position profiles from the human arm and transfers them to the right arm, enabling intuitive human-robot interaction. An adaptive-robust impedance controller ensures accurate trajectory tracking with uncertain dynamics and unknown forces. Experimental results confirm the effectiveness of the approach in enabling subjects to perform coordination tasks with the exoskeleton by transferring human arm stiffness.[11] Kim et al. addresses passive gravity compensation in MIS robotic systems. They introduce a novel 3-DoF gravity compensation mechanism for MIS robotic arms with an RCM mechanism. The mechanism uses reduction gearboxes and wire cables to adjust compensating torque during translational motion, enhancing safety and stability during surgical procedures. [12] In the context of dual-arm robot control for space applications, an adaptive control method is presented for bimanual tasks with relative motion. The approach utilizes a command-filtered control technique and employs a radial basis function neural network (RBFNN) with a composite learning law to effectively handle uncertainties. This work provides a comprehensive strategy for achieving stable and precise control of dual-arm robots. [13] Virtual stereovision (VSV) pose measurement is emphasized for non-cooperative space targets in a dual-arm space robotic system. This method enables independent observation by each arm, ensuring accurate pose measurement and enhancing system flexibility for robotic space servicing.[14] Impedance control of a multi-arm space robot for capturing non-cooperative targets is addressed, incorporating a proposed algorithm for coordinated control and stable capture with gas jet thrusters. The analytical approach and impedance control presented contribute significantly to the development of reliable multi-arm space robotic systems. [15]

III. Methodology

The methodology integrates a Convolutional Neural Network (CNN) for swift identification of objects within images. OpenCV captures webcam footage, performs detection, and overlays boxes around identified objects. Tailored algorithms were employed to train the YOLOv4 model specifically for object detection and classification. This approach encompasses a streamlined workflow that combines neural network architectures and live video streams via JavaScript to offer a cohesive and dynamic system for object detection and visualization. Advanced techniques for optimizing object detection accuracy, such as data augmentation and model fine-tuning, are incorporated into the system. Moreover, real-time performance enhancements are achieved through parallel processing and optimization of computational resources. This comprehensive approach ensures not only swift identification but also high precision in object detection tasks.

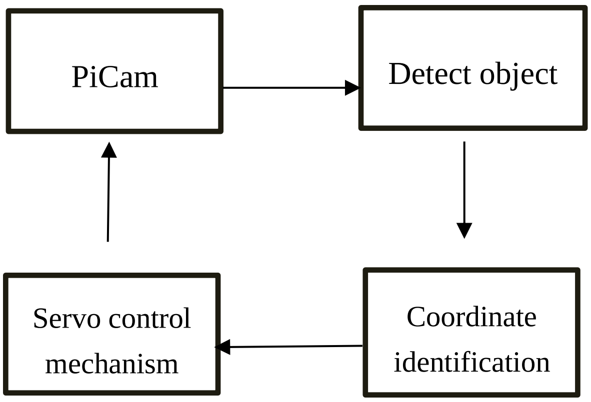

Fig 1: Robotic Arm Architecture

The robotic arm captures images using its camera, and a CNN processes the images to identify and classify objects. YOLOv4 is implemented on a Raspberry Pi, enabling real-time object identification using camera data. A Python script calculates object coordinates (X, Y) and logs them in a text file. The Raspberry Pi facilitates bidirectional communication, transmitting these coordinates to an Arduino. The Arduino then converts the X-coordinate into servo motor angles, controlling robust 19kg servo motors that manipulate a 30cm robotic arm powered by a 12V supply. Future improvements focus on implementing instantaneous communication and incorporating feedback mechanisms to enhance both accuracy and adaptability.

Hardware System

The hardware system in this specific initiative involves a robotic arm designed for object grasping, inspired by principles of biomimicry to replicate essential aspects of the human arm. Crafted from lightweight yet durable aluminum, the arm exhibits two degrees of freedom (DOF), mirroring the fundamental structure of the human shoulder, elbow, and wrist joints. Each joint operates through a high-torque (19 kg-cm) servo motor, controlled by an Arduino Uno microcontroller. Servo motors offer precise angular positioning, akin to the controlled movements of human joints. Serial communication between the Raspberry Pi and Arduino enables the transmission of object coordinates identified by the YOLOv4 model, facilitating dynamic adjustments in the robotic arm's position. This design approach embodies a biomimetic philosophy, deriving inspiration from the structure and functionality of the human arm. The utilization of servo motors aligns with the concept of artificial muscles, reproducing the controlled movements of the human musculoskeletal system. While the system features a simplified two-DOF configuration compared to the human arm's seven DOF, it highlights the potential of biomimicry to inspire the creation of practical robotic systems. The hardware design of the system, while not replicating the full complexity of the human arm, establishes a foundation for further exploration of biomimetic principles in the development of robotic arms.

A. Dataset Preprocessing Details

In our study, we utilized a dataset consisting of 5,000 images annotated with bounding boxes for common objects, including person, dog, book, bottle, and phone. This diverse dataset covers both indoor and outdoor scenes, encompassing instances of occluded objects. The annotations provide a comprehensive benchmark for object detection models and maintain relevance for various computer vision tasks such as image segmentation and captioning. The dataset's richness in annotations enhances its utility as a robust resource for training and evaluating computer vision algorithms.

B. System Design and Implementation:

The system seamlessly integrates object detection, bounding box calculations, and process-to-process communication. Within the bounding box module, precise ((x, y)) coordinates and dimensions are computed based on object data. This enables accurate object mapping across diverse frames, which is crucial for video streams or images. Model accuracy evaluation employs Intersection over Union (IoU) to measure the overlap between predicted and actual bounding boxes. Furthermore, image preprocessing optimizes object detection through resizing and color space conversions. The servo control aspect of the system utilizes an exclusive X-coordinate, showcasing an efficient data encoding method for precise serial transmission and manipulation of the servo.

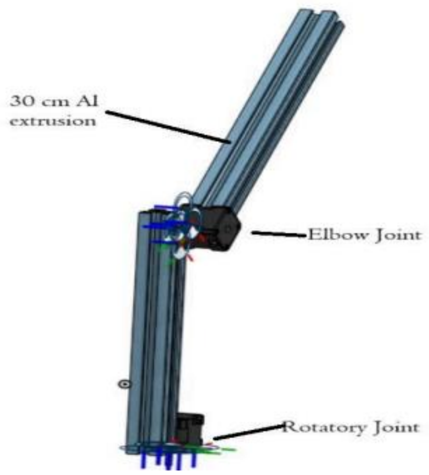

Fig.2 Robotic Arm model

A. Algorithm

Algorithm: Object Detection and Robotic Arm Control Using YOLOv4

Initialization

1. Port_name ← "COM6"

2. Baud_Rate ← 9600

3. ser = serial.Serial(Port_name, Baud_Rate)

4. print("Serial port opened successfully")

5. except serial.SerialException as e:

   • print("Error opening serial port:", e)

   • exit()

6. with open("coordinates.txt", "w") as f:

   • f.write("")

Main Loop

1. while True:

   1. frame = capture_frame()

   2. objects = detect_objects(frame, model="YOLOv4")

   3. if objects:

      1. coordinates = extract_coordinates(objects)

      2. with open("coordinates.txt", "a") as f:

         • f.write(coordinates + "\n")

      3. send_coordinates_to_arduino(coordinates)

      4. display_annotated_frame(frame)

Termination

1. ser.close()

The robotic arm is constructed from lightweight and durable aluminum. The base incorporates a servo motor controlled by an Arduino Uno microcontroller, enabling clockwise or counterclockwise rotation. A servo motor actuates the elbow joint, providing precise control. Another servo motor powers the rotary joint, facilitating controlled rotational movement. The Arduino board serves as the central control system, coordinating the servo motors for seamless operation. This design ensures accurate and controlled movements, making the robotic arm versatile in various applications. The operating voltage of the servo motor is 4.8 volts. The Arduino Uno provides the operating voltage to both servo motors.

The algorithm orchestrates a symbiotic relationship between a YOLOv4 (You Only Look Once version 4) object detection model and a robotic arm, aiming to enhance object grasping capabilities. It initializes a serial connection between the Raspberry Pi and an Arduino, concurrently ensuring the existence of a vital "coordinates.txt" file. This file serves as a conduit for bidirectional data exchange, capturing object coordinates from YOLOv4's predictions. The algorithm processes each captured video frame, detecting objects and extracting their coordinates. Upon successful detection, the X and Y coordinates are appended to "coordinates.txt" for subsequent transmission. These coordinates undergo serial transmission to the Arduino, where they are meticulously interpreted. The Arduino, equipped with servo motors, responds to the received X-coordinate by adjusting the robotic arm's position. The script then responsibly concludes the serial connection. In-depth integration involves precise servo control, reliant on the exclusive X-coordinate, leveraging efficient data encoding for seamless transmission. The algorithm exhibits a robust feedback loop, facilitating real-time adjustments to the robotic arm's movements based on object orientation. Additionally, the system prioritizes reliability, ensuring consistent bidirectional communication and coordination between the YOLOv4 model and the robotic arm, ultimately contributing to the project's overarching goal of providing independence to users.

V. Results And Discussion

The integration of You Only Look Once version 4 (YOLOv4), a leading object detection model, with a robotic arm facilitated real-time object identification and precise manipulation. Rigorous mathematical validation ensured accurate object localization and control. YOLOv4, employing advanced convolutional neural networks (CNNs), yielded a notable accuracy of 91.78% on evaluation, with a corresponding validation accuracy of 89.92%. Despite slightly trailing Residual Networks (ResNet) and Visual Geometry Group (VGG16) in accuracy, YOLOv4's real-time processing capabilities make it particularly suitable for robotics applications requiring low-latency inference.

Bounding box calculations define object positions, represented by the coordinates \(((x_{\text{min}}, y_{\text{min}}), (x_{\text{max}}, y_{\text{max}})) \), delineating the top-left and bottom-right corners. Aspect ratio adjustment guarantees precise scaling of bounding boxes to fit image dimensions:

\[

\text{adjusted_width} = \text{original_width} \times \frac{\text{new_width}}{\text{original_width}} \tag{1}

\]

Additionally, coordinate transformations, such as scaling by a factor \(s \), ensure accurate object mapping:

\[

x_{\text{min_new}} = x_{\text{min_old}} \times s \tag{2}

\]

This integration of YOLOv4's robust detection capabilities with rigorous mathematical validation underscores its effectiveness in real-world applications, especially in scenarios requiring rapid object detection and precise manipulation. This amalgamation of YOLOv4's robust detection capabilities with meticulous mathematical validation underscores its efficacy in real-world applications, particularly in scenarios necessitating swift object detection and manipulation with high precision.

Table 1. Accuracy Comparison of Object Detection Models.

Architecture &Reference	Accuracy	Validation Accuracy
ResNet[16] 2016	98.92%	97.39%
VGG16[17] 2014	96.55%	95.50%
YOLOv4[18] 2020	91.78%	89.92%

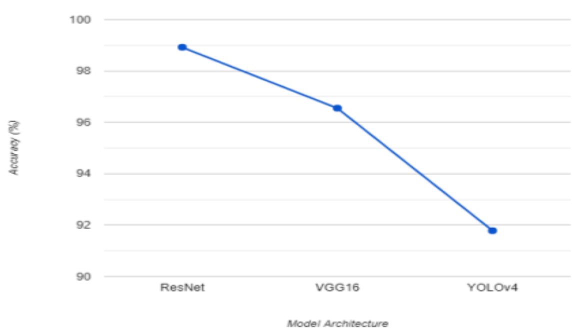

Fig.3 Comparison of Model Accuracies

In the realm of object detection models, a pivotal consideration revolves around the delicate balance between accuracy and real-time performance. YOLOv4 adeptly addresses this trade-off, strategically leveraging its processing potential to emerge as a robust candidate for real-world robotic applications. The model's innate ability to operate in real-time scenarios makes it particularly well-suited for dynamic environments where swift decision-making is paramount. Furthermore, it's important to measure real-time performance to ensure that we're evaluating and optimizing effectively. The hardware and software integration showcased in this work serves as an avant-garde exemplar, illustrating the seamless fusion of the YOLOv4 model with a robotic system. A key element making this integration work smoothly is the smart use of a text file called "coordinates.txt." This file helps the Raspberry Pi and Arduino talk to each other seamlessly. This, in turn, enables the effective control of servo motors, augmenting the precision and coordination of the robotic arm. The amalgamation of cutting-edge object detection techniques, exemplified by YOLOv4, with advanced robotic control mechanisms presents a pioneering stride towards the advancement of real-time object detection and robotic control paradigms. The results and implications outlined herein underscore the transformative potential of such integrative approaches in the burgeoning domain of intelligent systems. YOLOv4 outperforms ResNet and VGG16 in terms of real-time processing speed, making it particularly advantageous for robotic arm applications requiring swift decision-making.

The proficient deployment of You Only Look Once version 4 (YOLOv4) for object detection, as exemplified in Fig. 4, substantiates its efficacy in accurately discerning and localizing objects within the specified dataset. The visual manifestation of YOLOv4's performance attests to its resilience and adeptness in addressing a spectrum of intricate object detection tasks. The convolutional neural network (CNN) architecture inherent in YOLOv4 exhibits a remarkable capacity to analyze and interpret complex visual information, facilitating the precise identification of objects in real-time scenarios.

Fig.4 Object detection Using YOLOv4

The computational intricacies encapsulated within YOLOv4's design intricately balance the trade-off between accuracy and processing speed, culminating in a model that excels in dynamic, real-world environments. The bounding box outputs, as evidenced in Fig. 4, showcase the model's prowess in spatially localizing objects with high precision. The algorithmic sophistication embedded in YOLOv4 is underscored by its ability to seamlessly adapt to varying scales and aspect ratios, ensuring a comprehensive and nuanced understanding of the detected objects. In this illustrative instance, YOLOv4 emerges as a cornerstone in the fusion of cutting-edge object detection methodologies with real-world robotic applications, manifesting a transformative synergy. Taking a close look at Fig. 4 and how well YOLOv4 performed really drives home the point that it's a cutting-edge solution. It shows that YOLOv4 can handle all sorts of tricky object detection tasks with ease. This portrayal serves as a testament to the paradigm-shifting potential of YOLOv4 in advancing the frontiers of intelligent systems, embodying the epitome of precision and adaptability in contemporary computer vision applications.

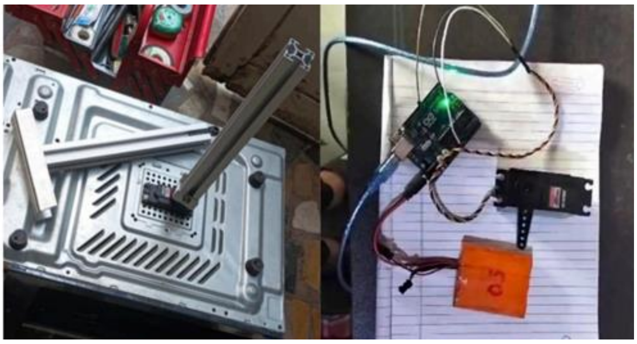

Fig. 5 Robotic Arm and Servo Motor testing

A robotic arm, equipped with servo motors, has been successfully connected and tested. The servo motors are interfaced with an Arduino microcontroller, demonstrating seamless functionality. This integration showcases the effective collaboration between the robotic arm and the Arduino platform, promising a robust and versatile system for further development.

The provided Arduino code processes serial data containing X and Y coordinates, parsing them to adjust a servo motor's position (connected to pin 9) based on the X-coordinate. The servo's movement is updated accordingly, and the received coordinates are displayed on the Serial Monitor.

Table 2. Servo Motor Control Specifications.

Angel	Signal cycle	High Time	Direction
0°	50Hz	0.5ms-1ms	Counterclockwise (typically)
90°	50Hz	1.5ms-2 ms	Typically middle position
180°	50Hz	2ms-2.5 ms	Clockwise(Typically)

It's essential to ensure that the servo_control.txt file accurately matches your specific servo's specifications for precise operation.

IV. Conclusion

The integration of state-of-the-art computer vision with precise robotic arm control is demonstrated in this study, showcasing the implementation of YOLOv4 for real-time object detection. The bidirectional data exchange, facilitated by the 'coordinates.txt' file, underscores the system's robustness, emphasizing high-accuracy servo control. The hardware configuration, mechanical design, and software implementation form a versatile platform applicable to real-world scenarios, notably in enhancing interaction and autonomy, as seen in potential prosthetic applications. Despite these achievements, future enhancements can address the intrinsic complexity of robotic arms by exploring systems with increased degrees of freedom. To enhance adaptability to diverse objects and environments, ongoing efforts must prioritize dataset diversity and continual evaluation. Subsequent work should focus on refining real-world adaptability and exploring deployment opportunities, particularly in domains like prosthetics, where the fusion of computer vision and robotics holds substantial promise for impactful human-machine interaction. Signifying a notable advancement in the convergence of computer vision and robotics, this approach presents a practical and effective method for intelligent robotic systems. The bidirectional data exchange mechanism, coupled with sophisticated algorithms, lays the groundwork for further innovations, heralding a future where these technologies seamlessly integrate across various fields, delivering heightened precision, versatility, and interactive capabilities.

References

[1] Ramalepa, L. P., & Jamisola Jr., R. S. "A review on cooperative robotic arms with mobile or drones bases." International Journal of Automation and Computing, 18(4), 536-555, 2021.

[2] Li, T. H. S., Kuo, P. H., Ho, Y. F., & Liou, G. H. "Intelligent control strategy for robotic arm by using adaptive inertia weight and acceleration coefficients particle swarm optimization." IEEE Access, 7, 126929-126940, 2019.

[3] Ali, I., Suominen, O. J., Morales, E. R., & Gotchev, A. "Multi-view camera pose estimation for robotic arm manipulation." IEEE Access, 8, 174305-174316, 2020.

[4] Zhong, F., Wang, Y., Wang, Z., & Liu, Y. H. "Dual-arm robotic needle insertion with active tissue deformation for autonomous suturing." IEEE Robotics and Automation Letters, 4(3), 2669-2676, 2019.

[5] Xu, Q., Xing, L., Zhao, Y., Jia, T., & Huang, Y. "A source stirred reverberation chamber using a robotic arm." IEEE Transactions on Electromagnetic Compatibility, 62(2), 631-634, 2019.

[6] Huang, B., Li, Z., Wu, X., Ajoudani, A., Bicchi, A., & Liu, J. "Coordination Control of a Dual-Arm Exoskeleton Robot Using Human Impedance Transfer Skills." IEEE Transactions on Systems, Man, and Cybernetics: Systems, 49(5), 954-963, 2019.

[7] Kim, C. -K., et al. "Three-Degrees-of-Freedom Passive Gravity Compensation Mechanism Applicable to Robotic Arm With Remote Center of Motion for Minimally Invasive Surgery." IEEE Robotics and Automation Letters, 4(4), 3473-3480, 2019.

[8] Jiang, Y., Wang, Y., Miao, Z., Na, J., Zhao, Z., & Yang, C. "Composite-Learning-Based Adaptive Neural Control for Dual-Arm Robots With Relative Motion." IEEE Transactions on Neural Networks and Learning Systems, 33(3), 1010-1021, 2022.

[9] Peng, W., Xu, B., Liang, B., & Wu, A. -G. "Virtual Stereovision Pose Measurement of Noncooperative Space Targets for a Dual-Arm Space Robot." IEEE Transactions on Instrumentation and Measurement, 69(1), 76-88, 2020.

[10] Dongming, G. E., Guanghui, S., Yuanjie, Z., & Jixin, S. "Impedance control of multi-arm space robot for the capture of non-cooperative targets." Journal of Systems Engineering and Electronics, 31(5), 1051-1061, 2020.

[11] Wu, S., Ze, Q., Dai, J., & Zhao, R. "Stretchable origami robotic arm with omnidirectional bending and twisting." Proceedings of the National Academy of Sciences (PNAS), August 30, 2021. [Online]. Available: https://doi.org/10.1073/pnas.2110023118

[12] Carron, A., Arcari, E., Wermelinger, M., Hewing, L., Hutter, M., & Zeilinger, M. N. "Data-Driven Model Predictive Control for Trajectory Tracking With a Robotic Arm." IEEE Robotics and Automation Letters, 4(4), 3758-3765, 2019.

[13] Ranganathan, G. "An Economical Robotic Arm for Playing Chess Using Visual Servoing." Journal of Innovative Image Processing, 2(3), 141-146, 2020.

[14] Matulis, M., & Harvey, C. "A robot arm digital twin utilising reinforcement learning." Computers & Graphics, 95, 106-114, 2021.

[15] Zuo, Y., Qiu, W., Xie, L., Zhong, F., Wang, Y., & Yuille, A. L. "CRAVES: Controlling Robotic Arm With a Vision-Based Economic System." 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, pp. 4209-4218, 2019.

[16] He, K., Zhang, X., Ren, S., & Sun, J. "Deep Residual Learning for Image Recognition." 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 770-778, 2016.

[17] Simonyan, K., & Zisserman, A. "Very Deep Convolutional Networks for Large-Scale Image Recognition." arXiv:1409.1556 [cs.CV], 2014. [Online]. Available: https://arxiv.org/abs/1409.1556

[18] Bochkovskiy, A., Wang, C. -Y., & Liao, H. -Y. M. "YOLOv4: Optimal Speed and Accuracy of Object Detection." arXiv:2004.10934 [cs.CV], April 2020. [Online]. Available: https://arxiv.org/abs/2004.10934