

## Analysis of Scene Perception for Human Robot Collaboration in Convolution Neural Network Environments

<sup>1</sup>Abdulkadir Hamidu Alkali, <sup>2</sup>Muhammad Zaharadeen Ahmed, <sup>2</sup>Abubakar Umar, <sup>2</sup>Mohammed Kabir Dauda, <sup>3</sup>Fatima Shittu, <sup>2</sup>Abdulsalam Ya'u Gital,  
<sup>1</sup>Department of Computer Engineering, University of Maiduguri, Borno State Nigeria  
<sup>2</sup>Department of Mathematical Sciences, Abubakar Tafawa Balewa University Bauchi, Bauchi State Nigeria  
<sup>3</sup>Department of Computer Science, Federal Polytechnic Damaturu, Yobe State Nigeria

### ABSTRACT

In this paper, analysis of robot visual reasoning (for pick and place) is conducted. This refers to reasoning the latent meaning of visual signals or indication for future robot actions from visual observations of an HRC scene. In this paper, projection matrix estimation is computed using 2D points to represent the figure plane and 3D points to represent scene detection for initiate pick and place operation. During simulation, equations are represented using matrix format based on figure coordinates. The design also implements automatic calibration and accuracy for robot workspace. In addition, vision-based management for the robot end effector is also conducted based on horizontal targets, and upright targets. However, pick and place is articulated in the experiment without visual control. In our results, analysis of joint tangential forces is computed for the three joints, Alink1, Blink2, and Clink3. The outcome of axial force variation of the three joints is conducted in their state of mobility. Finally, result of torque acting on the three joints increase when the simulation time is increased to achieve optimum performance.

### ARTICLE INFO

Article History  
 Received: July, 2022  
 Received in revised form: August, 2022  
 Accepted: September, 2022  
 Published online: September, 2022

### KEYWORDS

Mobility Matrix, Torque, Axial force, Tangential forces, and Effector

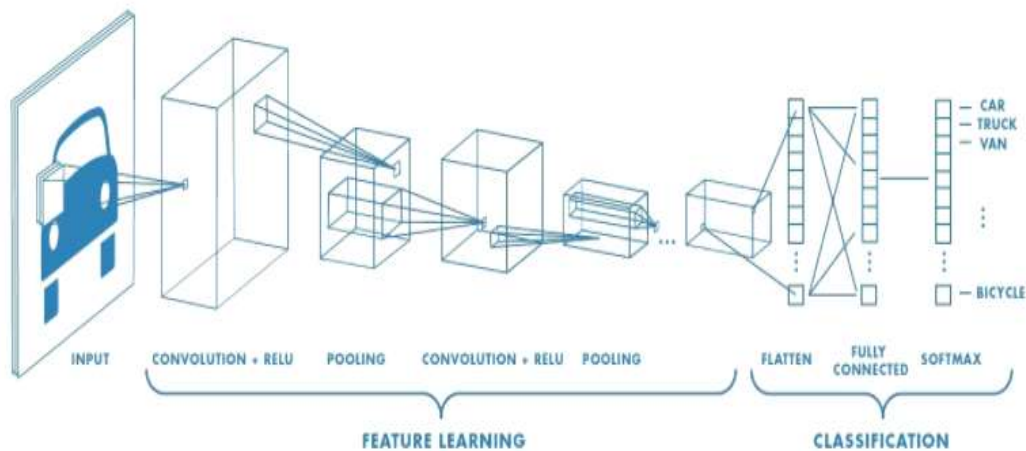


Figure 1: CNN Structure [11]

Human-robot collaboration (HRC) has gained attention in several areas of computer engineering, manufacturing, structures and

medical etc. This is due to the rapid and continuous increase in electronic devices, production efficiency and mass customization in

industries, homes, schools, and hospitals. HRC team combines the strength and accuracy of robots with the flexibility and inventiveness of people, allowing human operators and robots to work together in a shared workspace on shared tasks [4], [5].

## RELATED WORK

### A Convolution Layers

A convolutional layer is a set of parallel characteristics that are collected by sliding different convolution kernels. During each sliding position, a product operation results to a product and summation conducted between a convolution kernel and an input image. This process is carried out to project the information in the receptive field onto an element in the feature map. Convolution kernel's size is much smaller than the input image. Convolution layers consist of several characteristics. The first characteristic is local perception. The aim of using local perception is that researchers consider parts in the picture that are near to each other to be more relevant [6].

### B Robot Control Strategies

A major concern during implementation of industrial or academia HRC is its classification of physical contact. Research in [7] was conducted and concluded that the control approach cannot compensate for a poor mechanical design on their own. However, it remains a crucial characteristic in terms of performance, lessening sensitivity to indecision, and improvement of precision. The strategy of robot Control is classified into two [8], [12]. These are:

1. Pre-collision: The Pre-collision control strategies is centred on preventing damaging contact between the robot and its environment. This process involved the use of collision avoidance. This use sensory input such as camera, laser sensors etc to adjust the velocity of the body based on distance to independent quantity in the surrounding [9], [10].
2. Post-collision: The Post-collision strategies does not avoid possible contact between humans and robots. Instead, it lessens the contact force and the energy exchange between the two entities. This strategy is known as "interaction control strategies",

where the two famous sub-categories are termed direct and indirect control strategies. Direct control strategy is known as hybrid control strategies, whereas the indirect includes Admittance and Impedance control schemes. Direct-force control approaches control the manipulator's force along the constraint as well as the motion along with the directions of the unconstrained path by measuring the force [11], [12].

## METHODOLOGY

The robot has a functional camera device to visualize at 360 degrees. Calibration involves putting together both 3D and 2D points whereas calibrated Objects are mostly planes or cubes in two or three orthogonal forms. Chess sheet configuration are mostly included to develop the surfaces in equal distance and all through distributed precisely on the object exterior. The edges explain the well-known points within the robot coordinate system. Edge detection algorithm is used to identify the captured images. This ensures all the corners and edges are used in the process of projection matrix estimation. An alternative way to ensure this process is the self-calibration process. This approach computes the similarities from the camera adjustment when stationary.

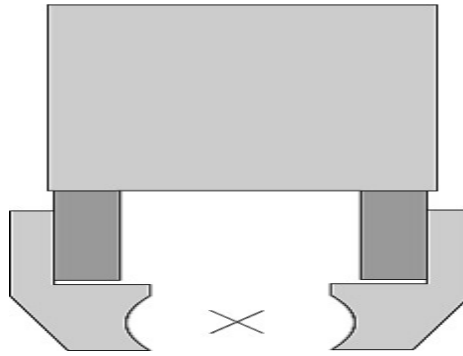
### Projection Matrix Estimation

The image plane surfaces represents the 2D points while scene of the situation of detection, pick and place represents the 3D points which are ascribed and can be deduced by the projection matrix, P as used in benchmark [1]:

The Elements in matrix P are set to be computed in terms of figure coordinates  $x_i$  and  $y_i$ . The associated 3D domain coordinates  $X_w$ ,  $Y_w$ , and  $Z_w$  as follows: A vector  $p = (p_{11}, \dots, p_{33})$  is defined, and equation (1) can be altered as  $Ap = 0$ . equation  $Ap=0$ , is defined by Matrix P having rank 11, requires at least  $N = 6$  correspondences.  $60 \times 12$  of matrix A is generated when  $N = 30$  computed equalities from the calibration site are used. Computation errors can easily be lessened using more than the required minimum of 6 calibration points. Separately from the trivial situation  $x = 0$ , no distinct solution of  $Ax = 0$  with the inclusion of signal noise. However, it is at

best to search for an exact answer, but one should lessen the computation cost function, such as the norm  $\|Ax\|$ . The norm  $\|UDVT x\|$  is limited using an outstanding value

decomposition, which factors matrix A into  $A = UDVT$ . The Single Value Decomposition (SVD) algorithm used in MATLAB simulation is used to invert Matrix A.



**Figure 1:** Gripper - Cross Hairs (Front View) [2]

Using the proposed approach, the camera adjustment and calibration can be structured based on the following:

1. Matrix approximation from 2D point intersection
2. Matrix Interpretation of the camera from the initial approximated matrix
3. Re-establishment of the Euclidean space using the identified control points from camera calibration (minimum of four non-coplanar control points).

Similarly, loss function is included in the proposed work to compute the degree of variation between the foreseen value and a real value. The value is a non-negative real value function. It is expressed as. Therefore, the standard form of the loss function can be expressed (as in equation 2) below.

The loss function is crucial in term of model optimization and enhancing precision. The proposed robot system is now more robust when the loss function is lesser. Also, the loss function remains a major part in empirical and structural risk function.

#### *Automatic Calibration and Accuracy for Robot Workspace*

As used in benchmark [2], experimental scenario in this study includes a Fancu M-16iB industrial robot, and two pairs of colored Charged Coupled Device sensors as camera are used between two ends of the workspace. Different capture of the robot used in benchmark [2] is presented (as in figure 2) holding a calibration slant at its geometrical point (cross hairs). The precision-made calibration tip fitted on the robot end effector can also be identified on same figure. The conical pointer of the calibration tip is painted red to make the color unique to the calibration tip and to make it easier to identify the tip in all calibration photographs. When coupled to the robot gripper, the tip was precision built to create an exact and repeatable reference. Through the known Denavit-Hardenberg geometry of the robot arm, the spatial tip coordinates can be traced back to the fixed robot base for any robot position. The Euclidean distance between a known 3D point  $P^0$  and its triangulated spatial representation  $P^t$  derived from the calibrated scene can be stated as follows:

$$\|P^0 - P^t\| = \sqrt{(P_x^0 - P_x^t)^2 + (P_y^0 - P_y^t)^2 + (P_z^0 - P_z^t)^2} \dots\dots\dots(3)$$

The precision is less towards centre of the scene of the two cameras. But the target is about equidistant from both. The precision of one observation increases as the calibration pointer moves away from center. Whereas the precision of the other decreases resulting to less precision.

*D. Vision Based management for Robot End Effector*

The process of vision control for pick operation consists of both horizontal and upright targets.

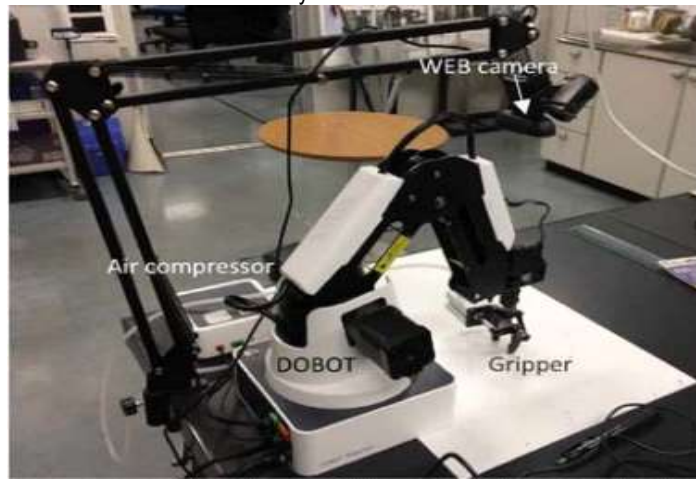
1. *Horizontal Targets:* The orientation of the object for Pick operation must be established to grasp horizontal cylindrical targets. The angle between the two extracted surface points and the world coordinate x axis may

be readily calculated to determine the orientation.

2. *Upright Targets:* During the target extraction phase of the robot, there is no need to distinguish between upright and horizontal items for the pick and place operation. The orientations of the objects have little influence on how the cylinder appears in the image.

*Pick and Place Without visual control*

A real pick and place experiment has been conducted using an existing test bed. The setup has been adopted from [3] (as in figure 2) below.



**Figure 2:** An Articulated experimental robot

Based on the robot position  $[x \ y \ \text{and} \ z]^T$  and the angle  $\theta$  of the gripper for the robot coordinate system is managed by a function as  $(x_a, y_a, z_a, \theta_a)$ . For the robot timer interrupt, a resolution of  $1600 \times 1200$  is captured. This is converted to binary in form of black and white. A connected device that has highest area is identified as the target object. The positions are represented as  $[l_x \ l_y]^T$  for  $(1 \leq l_x \leq 1600, 1 \leq l_y \leq 1200)$ . Therefore, the image coordinate system is then subtracted using the following equation below as follows.

Where as  $[X_1 \ Y_1]^T$  and  $[X_2 \ Y_2]^T$  of the robot coordinate system represent the different positions of left top and right down of the capture.

*Pick and Place with Visual Control*

Using this approach, camera configuration for visual management is not needed. This is not as in the case of without visual control. Based on (Nagata, F., et al, 2020), a lightweight endoscope camera has been placed near the gripper. Variable used during computation are varied as  $v(k) = [v_x(k) \ v_y(k)]^T$ . This is assisted for visualization of the robot and is generated using the following formula below.

Where  $k$  represents the discrete time.  $K_p$  and  $K_i$  represents the gains of the proportional and integral operation. Finally,  $e(k) = [e_x(k) \ e_y(k)]^T$  represent error section of the image coordinate

system. This is then computed by the formular below as.

Whereas  $X_d = [X_r/2 \ Y_{r+2}]^T$  and  $l(k) = [l_x(k) \ l_y(k)]^T$  represent the anticipated position and the computed object position from the image coordinate system. For the resolution of the system,  $[X_r \ Y_r]$  represent the captured image.

#### Designing the End Effector

The end effector is made up of three rotating components. Nearby linkages' axes are

usually perpendicular to one another. These are referred to as ALink 1, BLink 2, and CLink 3, based on their arrangement and working mechanism. The rotational direction of the axis is generally considered off from the end effector since Alink 1, Blink 2, and Clink 3 are rotated 180°, 90°, and 90°, respectively during mobile situation. In the simulation, the payload is a cubic shape with a side length of 250 mm and a mass of 100 g. All three links may spin 360 degrees. However, the ranges of the linkages are limited, (as shown in table 1) below.

**Table 1.** Selected ranges for joint movements

Link	Range
ALink 1	-180°C to 180°C
ALink 2	-90°C to 90°C
ALink 3	-90°C to 90°C

The workspace of the end effector might have two options depending on the initial location of Alink 3. In one situation, the workspaces of end points 1 and 2 are the same. These data points have two separate workspaces in the second example. To have the two options specified, the line connecting two end points must be perpendicular and parallel to the link's axis, respectively. The second circumstance is taken into account during the system simulation. When the workspace of the system is analysed, the end effector's axis of link 2 is parallel to the line connecting the two end points. For better mobility, the three links can spin one after the other or at the same time. To avoid unwanted vibration. ALink 1, Alink 2, and Alink 3 spin one after the other to avoid affecting precision and smooth motion. This represents the main function of this end-effector.

#### RESULT

This section presents the result of dynamic analysis of the proposed research. The magnitudes and directions of the forces and torques imparted to each component are determined using formulas and presentation using graphs. In the proposed research, the system is run with some predetermined inputs, and the forces and torques are displayed with respect to time. The beginning and ultimate positions of the end effector are ideal during their designs. The axial, Tangential forces and torques of the joints are represented as ALink 1, ALink 2, and ALink 3. These joints are enabled to rotates at angles 180°, 90°, and 45° respectively and the results can be depicted (as in figure 5) below.

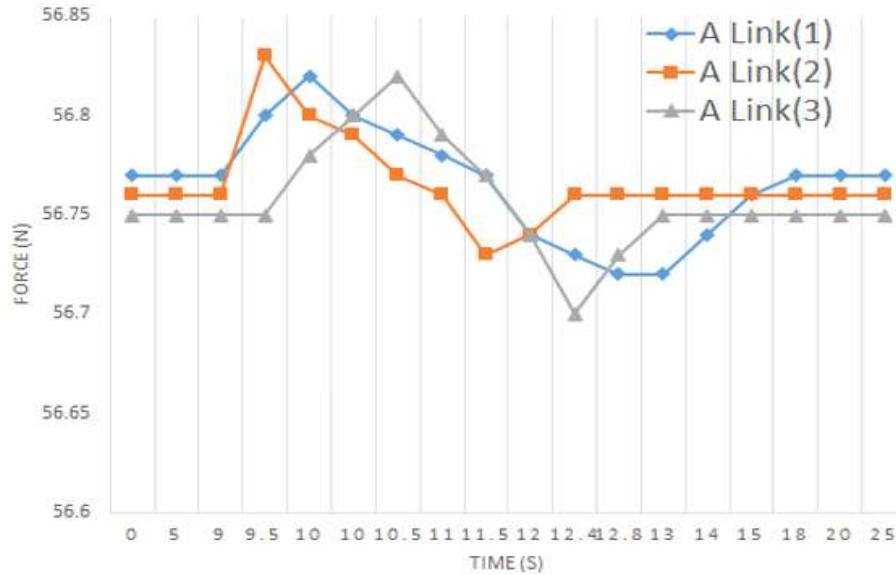


Figure 3: Joint Tangential forces

The simulation's payload is a cubic shape with a side length of 250mm and a mass of 100g. Angular acceleration and deceleration of the joints are kept constant during the simulation.

During the proposed design, when a joint attains a greatest angular velocity feasible, it begins to move at a constant angular velocity. The input angular displacements of the joints determine the shape of the delivered motion to the

joints. The following are the inputs given to the system during the simulation. Also, the component can execute an axial force - Figure 4, tangential force - Figure 3, and torque - Figure 6 on each joint. The magnitudes and orientations of the forces and torques based on the joints with respect to time can be represented in the graph (as in figure 4) below.

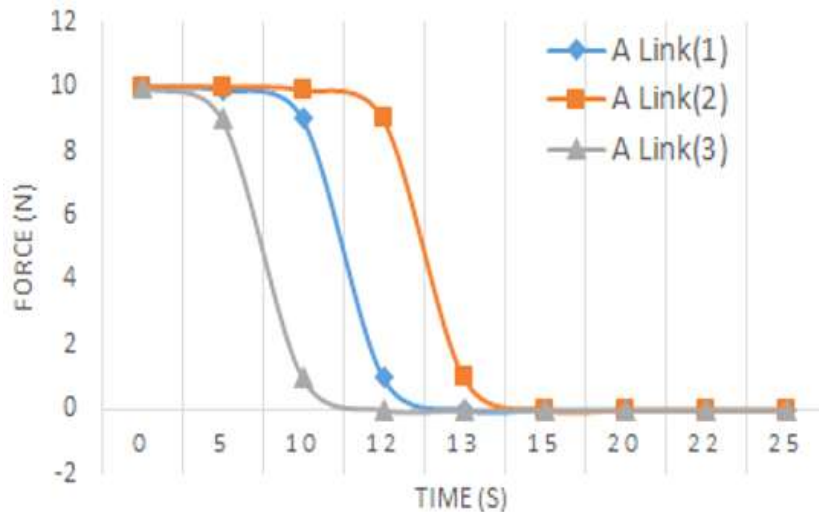


Figure 4: Axial force variation of three joints during mobility



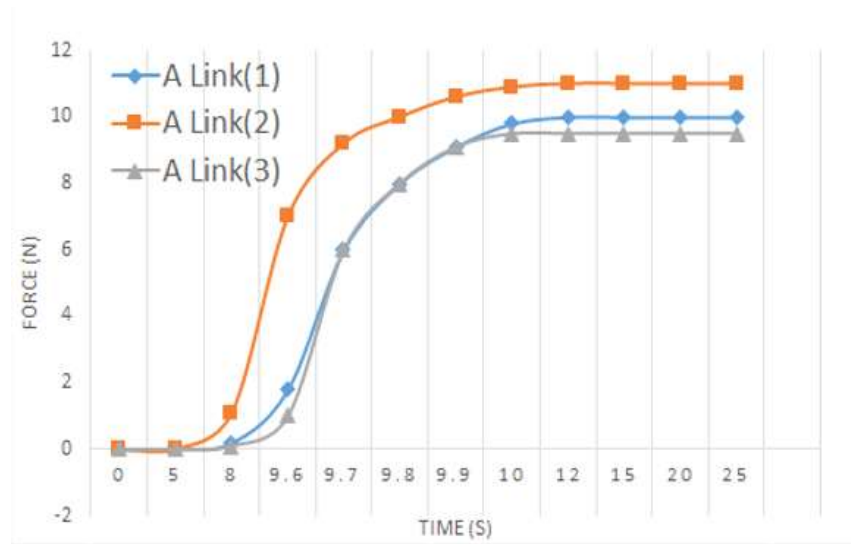


Figure 5: Joint Tangential forces

Based on mathematical computation, the position of the topmost i.e ALink3 is computed for an angular displacement for the three axes. By considering the coordinate on the plane  $[x_a, y_a, z_a]$  which is the base of the effector, and the robot is anchored and is achieved using isolation in the equation. It is important to note that all the three links are rotational. The computation is centered based on the strategy as follows.

For ALink 1  $0 \leq \theta_1 < 2\pi$ , where  $\theta_1$  represent the rotational

joint and as rotation angle. For ALink 2,  $-\pi/2 \leq \theta_2 < \pi/2$ ; where  $\theta_2$  represent the rotational joint and as rotation angle. For ALink 3,  $-\pi/2 \leq \theta_3 < \pi/2$ ; where  $\theta_3$  represent the rotational joint and as rotation angle (when  $\theta_2 = 90^\circ$  or  $270^\circ$ . Figure 8. below analyses the effector graphically further precisely for the three axes, where ALink 1 cannot rotate towards angle  $90^\circ$  or  $270^\circ$ .

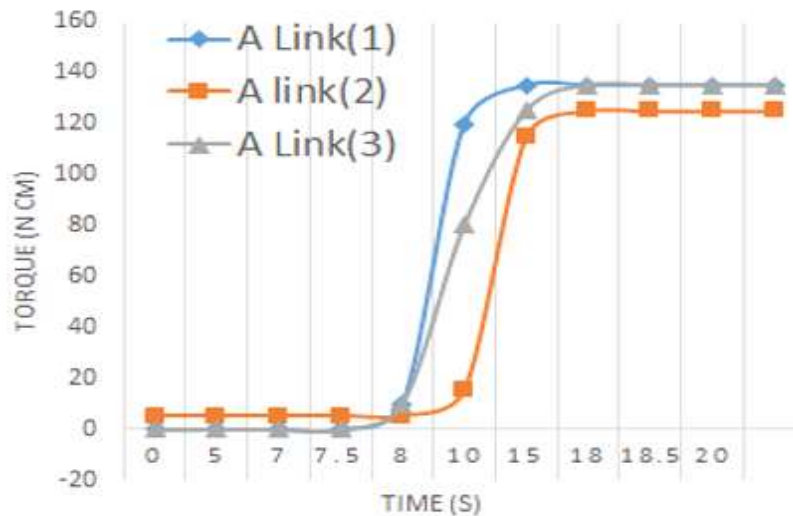


Figure 6. Torques acting on the joints during operation



## CONCLUSION

A Typical robotic machine for pick and place process must comprise of sensors for touch and vision. This must be attached to the robot at appropriate location for precise function. The paper conducted an experimental analysis that ensures functional operation and efficient strategies for reliable and flexible management of robots by means of automated stereo vision and touch systems. Projection Matrix Estimation and Automatic Calibration and Accuracy for Robot Workspace was conducted using equations for precision. This led to the creation of Vision Based Control of The Robot End Effector, Pick and Place Operations for both horizontal and upright targets.

## REFERENCE

- [1] Ansari, A. A., Akbar, S., & Sankaranarayanan, S. (2022). Precise Pick & Place System of 7-DoF KUKA Robot, Remote Control, Plotting & Computation of Trajectories, Cartesian Positions and Jacobian Matrices.
- [2] Cheng, C. Y., Renn, J. C., Chang, S. J., & Saputra, I. Development of a Simple Servo-Pneumatic 3-DOF Pick-and-Place Manipulator.
- [3] Shao, L., Nagata, F., Habib, M. K., & Watanabe, K. (2022). Visual Feedback Control Through Real-Time Movie Frames for Quadcopter with Object Count Function and Pick-and-Place Robot With Orientation Estimator. In *Handbook of Research on New Investigations in Artificial Life, AI, and Machine Learning* (pp. 99-116). IGI Global.
- [4] Chen, H., Kiyokawa, T., Wan, W., & Harada, K. (2022). Category-Association Based Similarity Matching for Novel Object Pick-and-Place Task. *IEEE Robotics and Automation Letters*.
- [5] Mey, J., Ebert, S., Lin, T., Nguyen, G. T., Gumhold, S., & Aßmann, U. (2022, January). Teaching Distributed and Heterogeneous Robotic Cells. In *2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC)* (pp. 1-2). IEEE.
- [6] Nagata, F., Miki, K., Otsuka, A., Yoshida, K., Watanabe, K., & Habib, M. K. (2020, October). Pick and place robot using visual feedback control and transfer learning-based CNN. In *2020 IEEE International Conference on Mechatronics and Automation (ICMA)* (pp. 850-855). IEEE.
- [7] Ansari, A. A., Akbar, S., & Sankaranarayanan, S. (2022). Precise Pick & Place System of 7-DoF KUKA Robot, Remote Control, Plotting & Computation of Trajectories, Cartesian Positions and Jacobian Matrices.
- [8] Smys, S., & Ranganathan, G. (2019). Robot assisted sensing control and manufacture in automobile industry. *Journal of ISMAC*, 1(03), 180-187.
- [9] Kazemi, S., & Kharrati, H. (2017). Visual processing and classification of items on moving conveyor with pick and place robot using PLC. *Intelligent industrial systems*, 3(1), 15-21.
- [10] Huang, P. C., & Mok, A. K. (2018, August). A case study of cyber-physical system design: Autonomous pick-and-place robot. In *2018 IEEE 24th international conference on embedded and real-time computing systems and applications (rtcsa)* (pp. 22-31). IEEE.
- [11] Sumit Saha (2018, December). A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way. <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>