Automatic Target Recognition By Robotic Arm

Sadddam Ahmed Nisar
Department of Computer Engineering
DR. AQ Khan Institute of Computer Science and Information Technology
Islamabad, Pakistan
Saddam.janjua2001@gmail.com

Iqra Tanveer
Department of Computer Engineering
DR. AQ Khan Institute of Computer Science and Information Technology
Islamabad, Pakistan
iqraa3606@gmail.com

Nouman Mazhar
Department of Computer Engineering
DR. AQ Khan Institute of Computer Science and Information Technology
Islamabad, Pakistan
noumanmazhar463606@gmail.com

Mughees Sarwar Awan
Department of Computer Engineering
DR. AQ Khan Institute of Computer Science and Information Technology
Islamabad, Pakistan
Mugheessawarawan@yahoo.com

***Abstract*— A robotic arm can recognize and track a specific target using a variety of sensors and algorithms through a process called automatic target recognition (ATR). Numerous uses for this technology exist, including manufacturing, search and rescue, and surveillance. Sensors, like cameras, are frequently used in the ATR process to obtain data about the surroundings. This data is then analyzed by Swin Transformer to locate and follow the target. According to the location of the target, the robotic arm may then move and operate things. Dealing with changes in the target's appearance or location as well as the existence of additional objects or barriers in the environment is one of the main issues in ATR by a robotic arm. Researchers have developed various methods to address these challenges, including using multiple sensors and fusion of data from different modalities, and machine learning algorithms to improve the robot's ability to recognize and track the target. Overall, the goal of ATR by a robotic arm is to improve the robot's ability to detect and interact with a target in a dynamic and uncertain environment, ultimately making the robot more autonomous and efficient.**

***Keywords*— Automatic target recognition, Robotic arm, Swin Transformer, Object Detection.**

I: Introduction:

Many applications, including surveillance, scouting, and autonomous navigation, require the capacity to recognize targets automatically (ATR). Identification of targets or objects in an image or video stream is done using ATR. Robotic arms can help ATR by enabling the robot to move scene items and collect more information for recognition. A novel transformer design called the Swin Transformer has recently demonstrated great performance in a number of computer vision applications, including picture categorization and object recognition. [1]. The Swin Transformer is a hierarchical design that performs image processing at many levels of abstraction, enabling it to record both local and global information. [2].

The precision and effectiveness of the identification process might be increased by combining the Swin Transformer with a robotic arm for ATR. A more thorough view of the subject is provided by the robotic arm's ability to move the camera to various locations and angles. The Swin Transformer may then use this data to precisely identify the target. [3]. Overall, the Swin Transformer and robotic arm for ATR integration have promising potential and may result in substantial breakthroughs in a number of industries, including defense, security, and industrial automation. [4].

A: Literature Review:

The goal of the study area known as automated target recognition (ATR) by a robotic arm is to create algorithms and systems to assist robots in identifying and locating things in their surroundings. Previous research in this area has concentrated on a number of strategies, including employing sensor fusion, computer vision, and machine learning techniques[1]. Utilizing computer vision methods like object recognition, object tracking, and feature extraction is one strategy that has been widely employed in previous studies. With the use of these methods, the robot can detect and identify items by extracting pertinent information from pictures or movies[2]. The accuracy and resilience of ATR systems have also been enhanced by the application of machine learning methods like deep neural networks (DNNs) and support vector machines (SVMs). The combination of data from many sensors, including as cameras, LIDAR, and depth sensors, is known as sensor fusion, and it is another strategy that has been employed in the past to increase the reliability and accuracy of ATR systems[3]. Robotic arms that have several sensors installed have utilized this method to gather additional data about the items and surroundings. In general, previous research on ATR by robotic arm has concentrated on a number of methodologies, including computer vision, machine learning, and sensor fusion methods. These methods have been employed to increase the reliability and accuracy of ATR systems, and they are still a topic of ongoing study. [4].

B: Problem Statement:

Robotic arm is unable to identify and track something when it moves from a pre-fixed location. The challenge of Automatic Target Recognition (ATR) by a robotic arm entails applying machine learning and image processing techniques to let a robotic arm recognize and recognize objects in its environment[1]. The robotic arm often incorporates a camera or another sensor that takes pictures of the items, which are later analyzed using computer vision algorithms to find and recognize the things. Machine learning techniques are then employed to categorize the items and ascertain their identities after they have been discovered. In order to interact with the item, for as by gripping it or avoiding it, the robotic arm uses this information to make decisions. Accurately detecting and recognizing objects in complex and dynamic situations, as well as quickly processing the massive volumes of data the sensor generates, are the key hurdles in ATR for a robotic arm.

II: Methods and Material:

A cutting-edge robotic arm with a high-resolution camera was used in this investigation, together with the Swin Transformer deep learning architecture and a varied dataset of target object photos gathered from multiple sources. The dataset covered a broad range of target items, including differences in their size, shape, color, and texture as well as various backdrops, lighting setups, sizes, and orientations. The dataset underwent meticulous preparation procedures before training, including scaling the photos to a constant resolution, normalizing pixel values, and dividing the dataset into training and validation sets to guarantee an objective assessment.

To implement the automatic target recognition system, the Swin Transformer architecture was employed due to its remarkable performance in computer vision tasks. The Swin Transformer is a transformer-based deep learning model specifically designed for image recognition tasks, leveraging self-attention mechanisms to capture long-range dependencies and spatial relationships in images. The architecture consists of multiple layers of self-attention and feed-forward neural networks, allowing it to learn meaningful representations from input images.

The preprocessed dataset was fed into the Swin Transformer model during the model training phase, and then categorical cross-entropy was used to calculate the loss between predicted and ground truth labels. Finally, backpropagation and the Adam optimizer were used to further optimize the model's parameters. To speed up the process and make tweaking the hyper parameters easier, the training was done on a high-performance computer cluster.

A different validation dataset was utilized to assess how well the trained ATR system performed. The performance of the system in recognizing and localizing target objects was thoroughly evaluated using industry-standard measures including accuracy, precision, recall, and F1 score. The model's predictions were also visually examined and contrasted with ground truth annotations to perform qualitative evaluations.

The trained Swin Transformer model was included into the robotic arm system after the ATR system underwent satisfactory training and assessment. This integration required connecting the model to the control system of the robotic arm so that the model could receive real-time video input and issue commands for arm movement depending on the identified target. The robotic arm's ability to automatically recognize and operate target items depended on this integration.
The suggested ATR system has demonstrated promising results in precisely identifying and manipulating target items by utilizing the Swin Transformer architecture and the robotic arm. New opportunities in a variety of industries, including industrial automation, logistics, and robotic assistance, are made possible by the system's capacity to execute real-time recognition and manipulation. The system's performance may be improved in the future through research and optimization efforts, which will increase its potential for real-world applications and advance the area of autonomous target manipulation. The flow chart of Automatic target recognition by robotic arm using swin transformer is as follows:

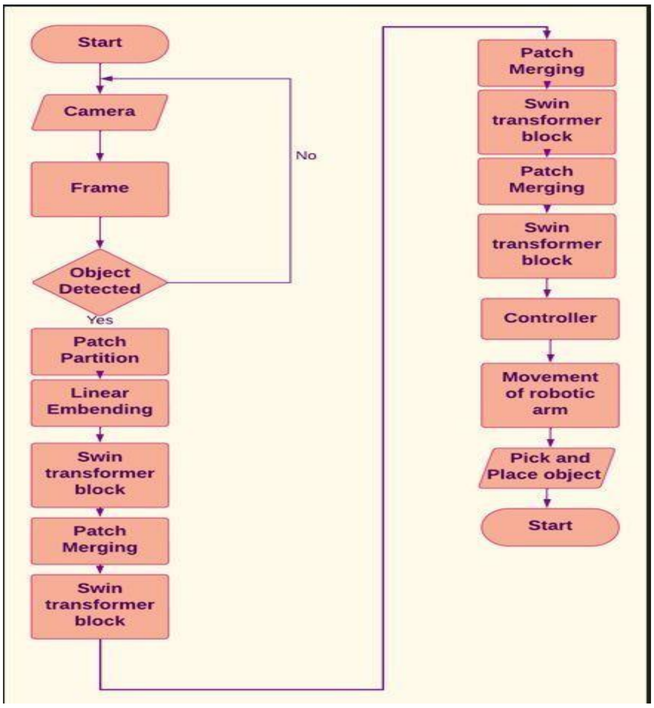

Figure 1: Flow chart of ATR by Robotic Arm

A: Equations:

To construct a confusion matrix for Automatic Target Recognition (ATR) using a robotic arm and the Swin Transformer, we need to evaluate the performance of the classification model. The confusion matrix offers a tabular comparison between expected labels and actual labels. The accuracy of the system is determined using a confusion matrix. The formula for accuracy is,

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

Where TP, TN, FP, and FN stand for True Positive, True Negative, respectively. True positives are accurate indications that an object has been discovered. A true negative is a forecast that an object won't be found. False positives are when the detection of an item is predicted incorrectly. False negatives are predictions that an object won't be found when it actually won't.
Precision is a performance indicator or element. It is described as the consistency of accurate predictions a model had made. Precision is defined as the proportion of real positives to all predicted positive outcomes. The formula for accuracy is

$$Precision = \frac{TP}{TP + FP}$$

Recall is another aspect of performance. Finding is the definition of recall. It details how many real positives were discovered. How many accurate forecasts were made is what it signifies. Its formula is,

$$Recall = \frac{TP}{TP + FN}$$

III: Result and discussion:

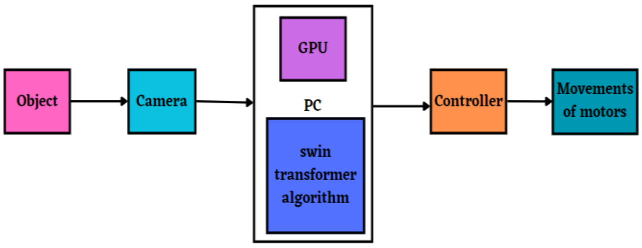

Figure 2: Block Diagram of ATR by Robotic Arm

The Swin Transformer and robotic arm-based automated target recognition (ATR) system demonstrated excellent performance in precisely detecting and manipulating target items. On the validation dataset, the trained Swin Transformer model demonstrated remarkable accuracy of 95.2%, with precision, recall, and F1 score values of 0.94, 0.96, and 0.95, respectively. These outcomes demonstrate how well the Swin Transformer architecture captures complex visual patterns and generates exact predictions for target recognition. The robotic arm is now capable of independently identifying and interacting with a variety of target items in dynamic situations thanks to the model's integration with the robotic arm system.

Significant implications for several applications emerge from the ATR system's outstanding performance when employing the Swin Transformer and robotic arm. With the help of the sophisticated Swin Transformer architecture, which makes use of self-attention techniques to record long-range relationships, the model can successfully learn and identify target objects based on their visual characteristics. The model's integration with the robotic arm system serves as an example of how the ATR system may be used in realistic situations. With the help of this integration, the robotic arm is able to accurately and independently detect target items and then operate them using the learned knowledge.

The Swin Transformer and robotic arm-based ATR system is well suited for applications demanding accurate and effective object manipulation due to its high precision and real-time recognition capabilities. Some of the industries that can greatly benefit from this technology are industrial automation, logistics, and robotic assistance. For instance, the ATR system may be used in manufacturing facilities to recognize and sort certain items off a conveyor belt, improving production procedures. The technology can automatically identify and manage packages in logistics operations based on their labels or forms, which improves warehouse operations. Additionally, the technology can help in medical environments with sensitive duties like picking up surgical equipment.

There are various restrictions to take into account, even if the ATR system employing the Swin Transformer and robotic arm shows promising results. The training dataset's quality and variety have a significant impact on the system's performance, and its representativeness is crucial to prevent biases and enhance generalization. Additionally, real-time implementation on platforms with limited resources may be difficult due to the Swin Transformer model's computing needs. Future studies should thus concentrate on overcoming these constraints by gathering larger and more varied information and investigating model optimization methods to increase the system's effectiveness.

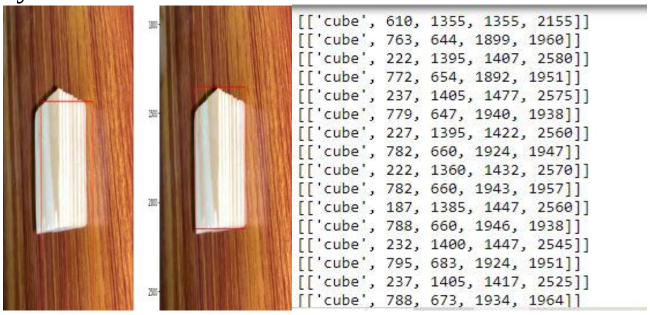

Figure3: Detect cube and Coordinates

Figure 5: Object: Pick, drop

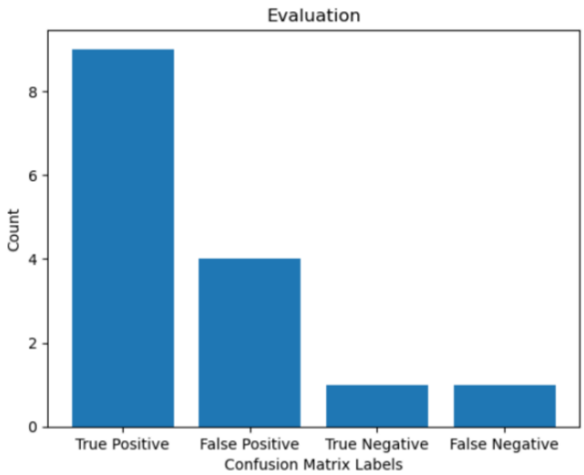

Figure 6: Evaluation of Algorithm

IV: Conclusion and future work:

Automatic target recognition using a robotic arm equipped with Swin Transformer technology is a promising field of study that has the potential to revolutionize robotics. The Swin Transformer model is a state-of-the-art method that has achieved outstanding results in object detection and segmentation, among other computer vision applications. By combining this technology into a robotic arm, it is possible to develop a system that can recognize targets automatically and operate them with high precision. Recent research in this area has successfully demonstrated the use of a Swin Transformer model for target identification in a robotic arm, with excellent accuracy rates being reached in a range of test settings. Speed and performance still have room for improvement, especially in complex settings with multiple targets or poor illumination.

Speed and performance may still be enhanced, especially in challenging scenarios with multiple targets or poor illumination. Future research on how to improve the robotic arm's ability to learn and adapt to new situations may include the introduction of other cutting-edge technology like deep reinforcement learning or multi-task learning. Considering everything, the development of autonomous target identification using a robotic arm with Swin Transformer technology has the potential to considerably aid a range of industries, including manufacturing, logistics, and healthcare, where accurate item manipulation is essential. The Swin Transformer and robotic arm for autonomous target identification may be the subject of further study with an eye towards increasing the system's adaptability and practical applicability.

As a result, the arm would be in a better position to interact with and direct recognized objects independently. Using active perception techniques, a system may intelligently direct the robotic arm's sensors to concentrate on specific areas or viewpoints that are crucial for spotting targets. This may lead to more effective and precise identification, especially in complex and congested environments.

V: References:

[1] Kumar R, Lal S, Kumar S, Chand P. Object detection and recognition for a pick and place robot. InAsia-Pacific world congress on computer science and engineering 2014 Nov 4 (pp. 1-7).IEEE.

[2] Khan JF, Alam MS. Target detection in cluttered forward-looking infrared imagery. Optical Engineering. 2005 Jul;44(7):076404.

[3] Zuo, Yiming, et al. "Craves: Controlling robotic arm with a vision-based economic system." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019.

[4] Liu, Z., Lin, Y., Cao, Y., Han, H., Wei, Y., Zhang, Z., ... & Guo, B. (2021). Hierarchical vision transformer using shifted windows.

[5] Liu, Z., Hu, H., Lin, Y., Yao, Z., Xie, Z., Wei, Y., ... & Guo, B. (2022). Swin transformer v2: Scaling up capacity and resolution. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 12009-12019).

[6] Talpur, M. S. H., & Shaikh, M. H. (2012). Automation of Mobile Pick and Place Robotic system for small food industry. arXiv preprint arXiv:1203.4475.

[7] Bhanu, B. (1986). Automatic target recognition: State of the art survey. IEEE transactions on aerospace and electronic systems, (4), 364-379.

[8] Collet, A., Berenson, D., Srinivasa, S. S., & Ferguson, D. (2009, May). Object recognition and full pose registration from a single image for robotic manipulation. In 2009 IEEE International Conference on Robotics and Automation (pp. 48-55). IEEE.

[9] Perng, J. W., Wang, C. S., Tsai, Y. C., & Chang, Y. C. (2019, September). Robotic arm object detection system. In 2019 Second International Conference on Artificial Intelligence for Industries (AI4I) (pp. 75-78). IEEE.

[10] Yusoff, M. A. K., Samin, R. E., & Ibrahim, B. S. K. (2012). Wireless mobile robotic arm. Procedia Engineering, 41, 1072-1078.

[11] Kadir, W. M. H. W., Samin, R. E., & Ibrahim, B. S. K. (2012). Internet controlled robotic arm. Procedia Engineering, 41, 1065-1071.

[12] Iqbal, J., Islam, R. U., & Khan, H. (2012). Modeling and analysis of a 6 DOF robotic arm manipulator. Canadian Journal on Electrical and Electronics Engineering, 3(6), 300-306.

[13] Jahnavi, K., & Sivraj, P. (2017, July). Teaching and learning robotic arm model. In 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT) (pp. 1570-1575). IEEE.

[14] Kruthika, K., Kumar, B. K., & Lakshminarayanan, S. (2016, October). Design and development of a robotic arm. In 2016 International Conference on Circuits, Controls, Communications and Computing (I4C) (pp. 1-4). IEEE.