



Algorithm for Object Collection Using a Robotic Manipulator and Machine Vision

Robinson Jiménez-Moreno, *PhD Ingeniería*¹, David García Veloza, *Estudiante Ingeniería Mecatrónica*², Javier Eduardo Martínez Baquero, *Msc. Tecnología Educativa y Medios Innovadores para la Educación*³

^{1,2}Universidad Militar Nueva Granada, Bogotá, Colombia, robinson.jimenez@unimilitar.edu.co,
est.david.garcia@unimilitar.edu.co



³Universidad de los Llanos, Villavicencio, Colombia, jmartinez@unillanos.edu.co

Abstract– *This document presents the detection of paper waste and its collection through a robotic agent in a mixed real-virtual validation system. An interface is designed in which the image of the robot's workspace is input, and the presence and location of paper waste are detected. For this purpose, a three-degree-of-freedom anthropomorphic robotic arm is implemented, and by calculating its kinematics, it moves to the waste location and collects it.*

Keywords-- *Object detection, kinematics, virtual environment, image processing, assistive robotics, manipulator robot.*

Algoritmo de Recolección de Objetos Mediante Manipulador Robótico y Visión de Máquina

Algorithm for Object Collection Using a Robotic Manipulator and Machine Vision

Robinson Jiménez-Moreno, *PhD Ingeniería*¹, David García Veloza, *Estudiante Ingeniería Mecatrónica*², Javier Eduardo Martínez Baquero, *Msc. Tecnología Educativa y Medios Innovadores para la Educación*³

^{1,2}Universidad Militar Nueva Granada, Bogotá, Colombia, *robinson.jimenez@unimilitar.edu.co*, *est.david.garcia@unimilitar.edu.co*

³Universidad de los Llanos, Villavicencio, Colombia, *jmartinez@unillanos.edu.co*

Resumen: El presente documento expone la detección de desechos de papel y su recolección mediante un agente robótico en un sistema de validación mixto real-virtual. Se diseña una interfaz en la que ingresa la imagen del área de trabajo del robot donde se detecta la presencia de desechos de papel y su localización mediante una red neuronal convolucional con detección por regiones de interés, para lo cual es implementado un brazo robótico antropomórfico de tres grados de libertad y mediante el cálculo de su cinemática, se desplaza a la ubicación del desecho y es recogido.

Palabras clave: Detección de objetos, cinemática, entorno virtual, procesamiento de imagen, robótica asistencial, robot manipulador.

Abstract– This document presents the detection of paper waste and its collection through a robotic agent in a mixed real-virtual validation system. An interface is designed in which the image of the robot's workspace is input, and the presence and location of paper waste are detected. For this purpose, a three-degree-of-freedom anthropomorphic robotic arm is implemented, and by calculating its kinematics, it moves to the waste location and collects it.

Keywords– Object detection, kinematics, virtual environment, image processing, assistive robotics, manipulator robot.

I. INTRODUCCIÓN

Los sistemas de visión por computador son empleados actualmente en aplicaciones para detección y seguimiento de objetos por medio de agentes robóticos que emplean algoritmos de inteligencia artificial (IA) [1]. Dentro de los algoritmos de IA que más destacan hoy día están los que

corresponden al aprendizaje profundo [2], los cuales emplean redes neuronales convolucionales (CNN) [3] para la detección de patrones en imágenes. De forma que al utilizar un robot con un sistema de visión e incluyendo un módulo RPN (Region Proposed Network) le permite identificar y localizar objetos en una escena [4].

El desarrollo de robots inteligentes capaces de interactuar con el medio que lo rodea implica esa capacidad de visión para identificación y localización, de forma que avances en reconocimiento de patrones basados en CNN que incluyen módulos RPN hoy día conocidas como Faster R-CNN, permiten implementar este tipo de robots [5][6]. Donde las aplicaciones más destacadas están en robótica orientada a recolección de fruta [7] y demás aplicaciones de recolección en cosechas [8].

Un método de implementación de redes convolucionales basadas en regiones es la transferencia de aprendizaje [9], donde se tiene arquitecturas definidas y bien conocidas en el estado del arte como la YOLO [10][11] y la ResNet en diferentes variaciones de profundidad en la arquitectura ResNet 50 y 101 [12]. El uso de transferencia de aprendizaje con estas arquitecturas reduce el tiempo de entrenamiento y de diseño de la red, presentado un alto desempeño en las tareas de identificación y localización en imágenes.

A su vez los agentes robóticos se han venido utilizando en la identificación y recolección de desperdicios, sea tipo basura [13] o específicos tipo desperdicios de papel [14]. Pero casos concretos como la recolección de desechos de papel en entornos residenciales no se encuentran en el estado del arte la

integración de la acción robótica con la identificación y localización, lo cual es el aporte del presente trabajo.

En la sección uno de este documento se expone el estado del arte y el aporte buscado, en la sección dos se exponen la metodología como el entrenamiento de la red y el brazo robótico de recolección, en la sección tres se exponen los resultados alcanzados y finalmente se exponen las conclusiones.

II. DESCRIPCIÓN DEL TRABAJO O METODOLOGÍA

El desarrollo de un robot manipulador asistencial que detecte un objeto en específico, como los desechos de papel, con el objetivo de despejar el área de trabajo, se basa en la visión por computadora que le permite al robot recoger el objeto. En la Figura 1 se ilustra el proceso llevado a cabo para el accionar del robot manipulador en el entorno virtual.

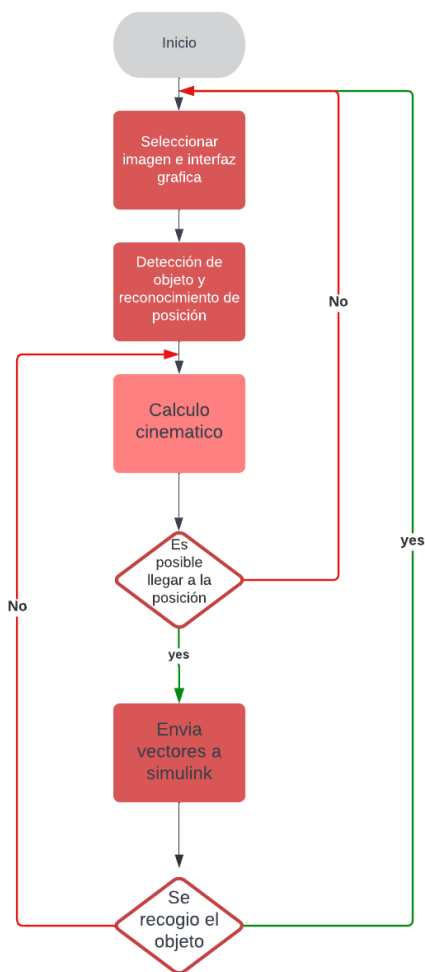


Fig. 1. Diagrama de flujo para accionar del robot manipulador en el entorno virtual

El algoritmo de identificación y ubicación se basa en una red del tipo Faster RCNN mediante transferencia de aprendizaje utilizando la arquitectura ResNet-50 [15][16], que requiere imágenes de tamaño 224*224 pixeles. La figura 2 ilustra la base de datos empleada donde se emplean variaciones de posición del papel y de fondo, también se agregan imágenes con más de un objeto para su identificación. El preprocesamiento empleado implica redimensionamiento de la imagen y adecuación de las dimensiones para conservar la uniformidad de la imagen de origen cuando no es cuadrada.



Fig. 2. Base de datos empleada

El entrenamiento de la red presenta una precisión del 91% y genera adecuadamente la caja que demarca la posición del objeto, como se aprecia en la Figura 3. Para el caso del entrenamiento se manejaron parámetros como un optimizador SGDM, 30 épocas de entrenamiento, una razón de aprendizaje de 1×10^{-3} , esta toma 46 minutos en equipo de cómputo con procesador Intel core I7, de 32 GB de memoria RAM y una GPU RTX3070 de 8 GB de memoria.



Fig. 3 Respuesta de la red para uno o dos objetos en la misma escena.

El algoritmo de control del brazo robótico fue implementado y validado en un entorno virtual respaldado por el entorno de desarrollo de MATLAB®. La Figura 4 ilustra este entorno. Los bloques de la Figura 1 se desarrollan en este entorno. Para la detección del objeto, en la interfaz gráfica se encuentran dos botones: uno para seleccionar la imagen y otro para la detección del objeto, de esta manera se permite conocer su posición.

Para el movimiento del robot, es necesario calcular la cinemática inversa, directa y diferencial del robot. Esto se realiza dentro de un ciclo para calcular iterativamente los cambios en el manipulador. Se comienza realizando el método Denavit-Hartenberg (D-H) [17] para la posición de home que se visualiza en la Tabla 1.

TABLA 1
MÉTODO DENAVIT-HARTENBERG

I	a_i	α_i	d_i	θ_i
1	0	α_1	d_1	θ_1
2	a_1	0	0	θ_2
3	a_2	0	0	θ_3

Para comenzar, se realiza el cálculo de la cinemática inversa del robot mediante un análisis matemático que evalúa su posición final.

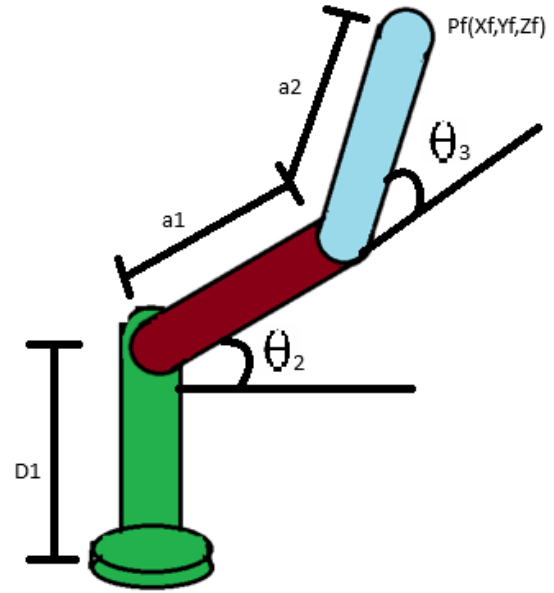


Fig. 4. Ambiente virtual de pruebas

Para comenzar, se calcula la distancia desde el punto inicial hasta el punto final utilizando la raíz de la suma de las diferencias al cuadrado. Luego, para encontrar el coseno de theta 3 mediante la ley de los cosenos, se resta la distancia d1 y se divide por el doble del producto de a1 y a2 como se muestra en la ecuación (1).

$$\cos(\theta_3) = \frac{(X_f^2 + Y_f^2 + (z_f - d_1)^2 - a_1^2 + a_2^2)}{2 * a_1 * a_2} \quad (1)$$

Una vez que se conoce el coseno de theta 3, se obtiene el valor del seno de theta 3 mediante identidades trigonométricas como se observa en la ecuación (2).

$$\sin(\theta_3) = \sqrt{1 - \cos(\theta_3)^2} \quad (2)$$

Conociendo estos dos valores, se utiliza la función atan2d para determinar el valor del ángulo. Esta misma función se utiliza para el valor de theta 1, utilizando los valores correspondientes de X y Y, como se observa en las ecuaciones (3) y (4).

$$\theta_3 = \text{atan2d}(\sin(\theta_3), \cos(\theta_3)) \quad (3)$$

$$\theta_1 = \text{atan2d}(X_f, Y_v) \quad (4)$$

La ecuación (5) permite determinar el valor de theta2 es igual a la diferencia entre dos valores arco tangentes (atan2). El primer atan2

involucra la diferencia entre la coordenada z del extremo del efector (Zf) y la longitud d1, dividida por la raíz cuadrada de la suma de las coordenadas Xf al cuadrado y las coordenadas Yf al cuadrado. El segundo atan2 implica el producto de a3 por el seno de theta3, dividido por la suma de a2 y el producto de a3 por el coseno de theta3.

$$\theta_2 = \text{atan2d} \left(Zf - d1, \sqrt{Xf^2 + Yf^2} \right) - \text{atan2}(a2 * \sin(\theta_3), a1 + a2 * \cos(\theta_3)) \quad (5)$$

Una vez que se conocen los valores de theta 1, theta 2 y theta 3, se calculan los valores de las matrices de transformación, como se muestra en las ecuaciones (6), (7) y (8).

$$a_1 \begin{bmatrix} \cos(\theta_1) & -\sin(\theta_1) * \cos(\alpha_1) & \sin(\theta_1) * \sin(\alpha_1) & a0 * \cos(\theta_1) \\ \sin(\theta_1) & \cos(\theta_1) * \cos(\alpha_1) & \sin(\theta_1) * \sin(\alpha_1) & a0 * \sin(\theta_1) \\ 0 & \sin(\alpha_1) & \cos(\alpha_1) & d1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

$$a_2 \begin{bmatrix} \cos(\theta_2) & -\sin(\theta_2) * \cos(\alpha_2) & \sin(\theta_2) * \sin(\alpha_2) & a1 * \cos(\theta_2) \\ \sin(\theta_2) & \cos(\theta_2) * \cos(\alpha_2) & \sin(\theta_2) * \sin(\alpha_2) & a1 * \sin(\theta_2) \\ 0 & \sin(\alpha_2) & \cos(\alpha_2) & d2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (7)$$

$$a_3 \begin{bmatrix} \cos(\theta_3) & -\sin(\theta_3) * \cos(\alpha_3) & \sin(\theta_3) * \sin(\alpha_3) & a2 * \cos(\theta_3) \\ \sin(\theta_3) & \cos(\theta_3) * \cos(\alpha_3) & \sin(\theta_3) * \sin(\alpha_3) & a2 * \sin(\theta_3) \\ 0 & \sin(\alpha_3) & \cos(\alpha_3) & d3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (8)$$

Una vez que se tienen las matrices de transformación, el siguiente paso es derivar las relaciones que relacionan las coordenadas del extremo del robot con las variables de articulación. Esto se hace utilizando el producto de las matrices de transformación. Para cada articulación, se obtienen las coordenadas del extremo del robot en términos de las variables de articulación correspondientes.

Después de obtener estas relaciones, se procede a calcular el jacobiano [18][19]. El jacobiano es una matriz que relaciona las velocidades de las variables de articulación con las velocidades de las coordenadas del extremo del robot. Para hacer esto, se calcula el producto cruz entre los vectores de posición generados por las matrices de transformación y se organizan en una matriz que representa el jacobiano. El jacobiano resultante tendrá tres filas (una por cada coordenada del extremo del robot) y tres columnas (una por cada variable de articulación).

III. RESULTADOS Y DISCUSIÓN

Las pruebas de recolección se realizan mediante el uso del jacobiano, que aplicado a la cinemática de robots permite analizar cómo pequeños cambios en las articulaciones afectan la posición y orientación del extremo del robot. El proceso para hallar el jacobiano del robot RRR de 3 grados de libertad [20] empleado involucra la derivación de relaciones entre las variables de articulación y las coordenadas del extremo del

robot a través de las matrices de transformación, seguido por el cálculo del jacobiano mediante el producto cruz de los vectores de posición.

La posición o coordenadas de desplazamiento del efector final del robot se toman en función al centroide del recuadro de la detección del papel, este es generado por la red tipo Faster RCNN. Como se aprecia en la figura 3 los niveles de confianza en la detección de papel son superiores al 90%, lo cual permite validar el punto de recolección sobre el objeto.

La figura 5 ilustra la aplicación (App) en MATLAB desarrollada para validar las pruebas, en esta se evidencia el modelo D-H del robot, un botón para cargue de la imagen del área de detección del robot, sobre la que se detecta el centroide, para determinar el punto de agarre por el efector. Se destacan las posiciones del objeto en coordenadas x,y , así como la posición final del efector empleando las matrices de transformación expuestas en el App. En función a estos valores se determinarán los porcentajes de error en la labor de recolección del robot. El recuadro final de la posición angular por articulación permite validar la simulación en el desplazamiento desde la posición de reposo o home del robot al de recogida del papel.

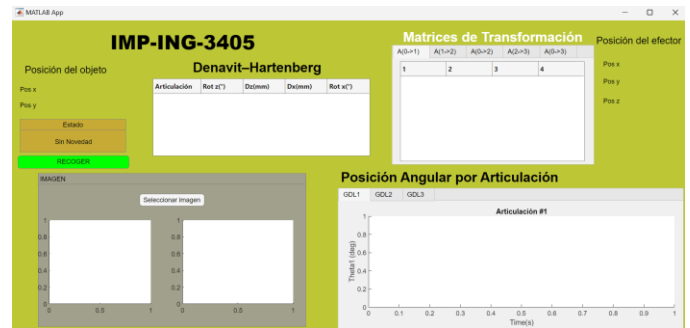


Fig. 5. Ambiente virtual de pruebas.

A continuación, en las figuras 6, 7, 8 y 9 se exponen los resultados de identificación de posición del papel (objeto) y del posicionamiento del efector final del robot para su agarre.



Fig. 6. Ambiente virtual prueba 1.

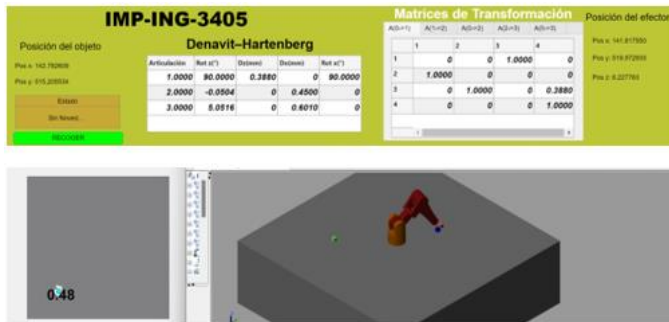


Fig. 7. Ambiente virtual prueba 2.

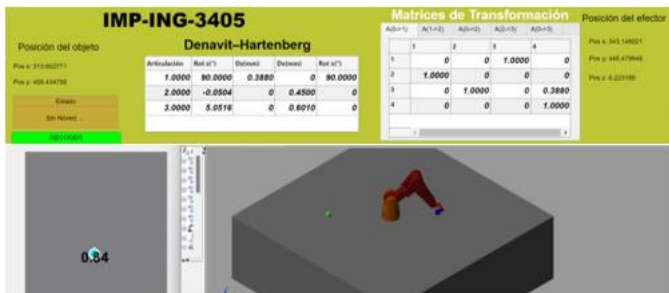


Fig. 8. Ambiente virtual prueba 3.

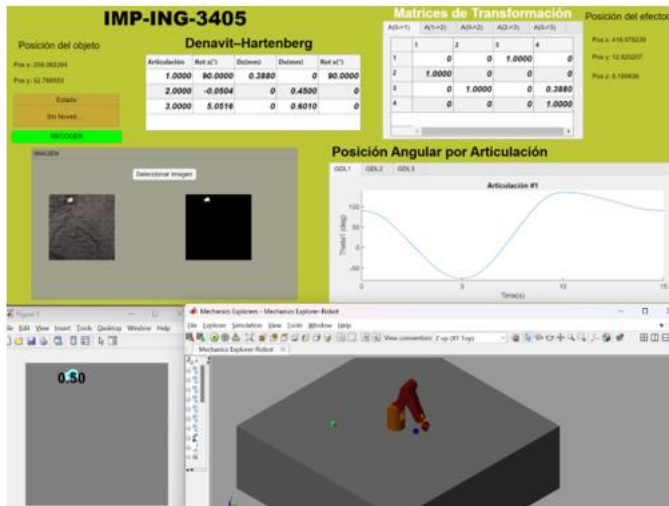


Fig. 9. Ambiente virtual prueba 4.

Las pruebas anteriores dan por resultado la medición del error en la desviación del efector con relación al papel dado en milímetros, como se exponen en la Tabla 2.

TABLA 2
MEDICIÓN DEL ERROR DE POSICIÓN

Prueba	Posición papel	Posición efector	Error
1	X:584 Y:359	X:637 Y:310	22.9
2	X:142 Y:515	X:141 Y:519	3.59
Z	X:313 Y:458	X:343 Y:445	7.11
4	X:209 Y:52	X:416 Y:12	200.8

Se aprecia que el error aumenta hacia la frontera del efector dada las limitaciones del área de trabajo e imposibilidad del efector de llegar a dicho lugar, como se evidencia en las pruebas 1 y 4, siendo más notoria en la prueba 4 debido a que el robot no puede acercarse más a su eje.

IV. CONCLUSIONES

Se logró validar mediante el reconocimiento de desechos de papel, la manipulación correcta de los mismos por parte del brazo robótico. Logrando evidenciar que desde que el desecho se encuentre en el campo de visión de la cámara que soporta el sistema de percepción del entorno y a su vez en el campo de trabajo del robot, el robot toma acertadamente el desecho y lo deposita en el punto final deseado.

Se toman como referencia de coordenadas la base del robot en relación con las coordenadas de la cámara, para coordinar el punto de localización y el del efector final del robot. Como trabajo futuro se encuentra el validar el esquema realizando navegación de un robot móvil que soporte el brazo y explore un entorno amplio de trabajo para la búsqueda de los desechos, e implementar un controlador de posición para obtener error cero entre las ubicaciones del objeto y el efector.

La detección de la red de aprendizaje profundo empleada entrega coordenadas de posicionamiento en relación a la ubicación de la cámara, e independiente del área que cubra el tamaño del papel empleado, sin embargo, trozos pequeños de radio inferior a 5 milímetros no son agarrados por el brazo robótico real en que se basó la simulación, a lo cual debe plantearse un ajuste de la forma del gripper.

AGRADECIMIENTOS

Los autores agradecen a la vicerrectoría de investigaciones de la Universidad Militar Nueva Granada, por el financiamiento de este proyecto con código IMP-ING-3405 (2021-2022) y titulado "Prototipo robótico móvil para tareas asistenciales en entornos residenciales". A su vez agradecen a

la Universidad de los Llanos por su participación en el desarrollo del presente proyecto.

REFERENCIAS

- [1] A. Kamble, A. C. Mitra, A. Tathe, S. Kumbharkar, and A. Bhandare, "Object following robot based on AI/ML," *Materials Today: Proceedings*, vol. 72, no. 3, pp. 1817-1824, 2023. DOI: 10.1016/j.matpr.2022.09.577.
- [2] Y. Zhang, M. Li, M. Zhang, C. Guo, and Z. Jiang, "Fast inspection and accurate recognition of target objects for astronaut robots through deep learning," *Measurement*, vol. 213, 2023, article 112687. DOI: 10.1016/j.measurement.2023.112687.
- [3] T.-W. Wu, H. Zhang, W. Peng, F. Lü, and P.-J. He, "Applications of convolutional neural networks for intelligent waste identification and recycling: A review," *Resources, Conservation and Recycling*, vol. 190, 2023, article 106813. DOI: 10.1016/j.resconrec.2022.106813.
- [4] T. K. Shashank, N. Hitesh, and H. S. Gururaja, "Application of few-shot object detection in robotic perception," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 114-118, 2022. DOI: 10.1016/j.gtp.2022.04.024.
- [5] Y. Yu, K. Zhang, L. Yang, and D. Zhang, "Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN," *Computers and Electronics in Agriculture*, vol. 163, 2019, article 104846. DOI: 10.1016/j.compag.2019.06.001.
- [6] X. Chen and J. Guhl, "Industrial Robot Control with Object Recognition based on Deep Learning," *Procedia CIRP*, vol. 76, pp. 149-154, 2018. DOI: 10.1016/j.procir.2018.01.021.
- [7] Y. Yang, Y. Han, S. Li, Y. Yang, M. Zhang, and H. Li, "Vision based fruit recognition and positioning technology for harvesting robots," *Computers and Electronics in Agriculture*, vol. 213, 2023, article 108258. DOI: 10.1016/j.compag.2023.108258.
- [8] S. Wan and S. Goudos, "Faster R-CNN for multi-class fruit detection using a robotic vision system," *Computer Networks*, vol. 168, 2020, article 107036. DOI: 10.1016/j.comnet.2019.107036.
- [9] S. Hou, B. Dong, H. Wang, and G. Wu, "Inspection of surface defects on stay cables using a robot and transfer learning," *Automation in Construction*, vol. 119, 2020, article 103382. DOI: 10.1016/j.autcon.2020.103382.
- [10] H. Gharakhani, J. A. Thomasson, and Y. Lu, "Integration and preliminary evaluation of a robotic cotton harvester prototype," *Computers and Electronics in Agriculture*, vol. 211, 2023, article 107943. DOI: 10.1016/j.compag.2023.107943.
- [11] S. Hu, A. Feng, J. Shi, J. Li, F. Khan, H. Zhu, J. Chen, and G. Chen, "Underwater gas leak detection using an autonomous underwater vehicle (robotic fish)," *Process Safety and Environmental Protection*, vol. 167, 2022, pp. 89-96. DOI: 10.1016/j.psep.2022.09.002.
- [12] Y.-L. Zhang, Y.-C. Kim, and G.-W. Cha, "Assessment of deep learning-based image analysis for disaster waste identification," *Journal of Cleaner Production*, 2023, article 139351. DOI: 10.1016/j.jclepro.2023.139351.
- [13] R. Aarathi, G. Rishma, "A Vision Based Approach to Localize Waste Objects and Geometric Features Extraction for Robotic Manipulation," *Procedia Computer Science*, vol. 218, 2023, pp. 1342-1352. DOI: 10.1016/j.procs.2023.01.113.
- [14] K. Pivnenko, E. Eriksson, and T. F. Astrup, "Waste paper for recycling: Overview and identification of potentially critical substances," *Waste Management*, vol. 45, 2015, pp. 134-142. DOI: 10.1016/j.wasman.2015.02.028.
- [15] L. Vuyyuru, S. Nalluri, J. Vempatapu and R. b. Thopuri, "Air Pollution Detection using Resnet-50," 2023 7th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2023, pp. 845-849, doi: 10.1109/ICCMC56507.2023.10083837.
- [16] Z. Zahisham, C. P. Lee and K. M. Lim, "Food Recognition with ResNet-50," 2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), Kota Kinabalu, Malaysia, 2020, pp. 1-5, doi: 10.1109/IICAIET49801.2020.9257825.
- [17] O. Hock, P. Drgoňa and M. Paškala, "Simulation model of adjustable arm using Denavit-Hartenberg parameters," 2014 ELEKTRO, Rajecke Teplice, Slovakia, 2014, pp. 176-179, doi: 10.1109/ELEKTRO.2014.6847896.
- [18] X. Deng, C. Liu, W. Li, M. Tong, X. Yu and W. Lin, "Camera Intrinsic Invariance of Image Jacobian in 4 DOF Image Based Visual Servo," IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society, Singapore, 2020, pp. 4750-4755, doi: 10.1109/IECON43393.2020.9254376.
- [19] C. C. Cheah, C. Liu and H. C. Liaw, "Stability of inverse Jacobian control for robot manipulator," Proceedings of the 2004 IEEE International Conference on Control Applications, 2004., Taipei, Taiwan, 2004, pp. 321-326 Vol.1, doi: 10.1109/CCA.2004.1387231.
- [20] Guilin Yang, Weihai Chen and I-Ming Chen, "A geometrical method for the singularity analysis of 3-RRR planar parallel robots with different actuation schemes," IEEE/RSJ International Conference on Intelligent Robots and Systems, Lausanne, Switzerland, 2002, pp. 2055-2060 vol.3, doi: 10.1109/IRDS.2002.1041568.