# Integrating LLMs into Collaborative Robotics for Automated Zipped Apparel Disassembly


1st Andrea Bonci
*Dept. of Information Engineering*
*Polytechnic University of Marche*
Ancona, Italy
a.bonci@univpm.it

2nd Alessandro Di Biase
*Dept. of Electrical and Information Engineering*
*Polytechnic of Bari*
Bari, Italy
a.dibiase8@phd.poliba.it

3rd Sauro Longhi
*Dept. of Information Engineering*
*Polytechnic University of Marche*
Ancona, Italy
sauro.longhi@univpm.it

4th Ilaria Pellicani
*Dept. of Information Engineering*
*Polytechnic University of Marche*
Ancona, Italy
i.pellicani@staff.univpm.it

5th Mariorosario Prist
*Dept. of Information Engineering*
*Polytechnic University of Marche*
Ancona, Italy
m.prist@staff.univpm.it

6th Andrea Serafini
*Dept. of Information Engineering*
*Polytechnic University of Marche*
Ancona, Italy
andrea.serafini@pm.univpm.it



***Abstract*—In order to successfully integrate circular economy principles into current value chains, it is crucial to ensure the economic sustainability of disassembly processes. Recent scientific and technological innovations, including Large Language Models (LLMs) in artificial intelligence (AI), are greatly accelerating progress in enhancing robotic capabilities. Dismantling is the first step in the re-manufacturing, repair, and recycling processes of end-of-life (EoL) products. Traditionally, this operation is performed manually by operators or by expensive dedicated robotic cells. Although manual disassembly offers flexibility in handling complex situations, it is a time-consuming and labour-intensive process that can negatively affect the health of operators and the cost-effectiveness of the disassembly process. However, robot disassembly presents difficulties in handling complex parts in a flexible manner. Robot application emerges as an automated versatile solution, capable of handling uncertainties in the frequency, quantity, and quality of end-of-life products. A fully automated approach for the removal of clothing zips from garments is proposed. The approach detects the pixels corresponding to the zip using LLM-Based AI techniques to plan the path of the robotic arm. The solution reduces the effort of AI training in industrial applications. Simulations validate the approach and its performance.**



***Index Terms*—Disassembly, Collaborative Robotics, Large Language Models, Artificial Intelligence, Apparel, Clothing, EoL Products, Circular Economy.**

The work was co-funded by
- "CIRCULAR FASHION - IN.CO. S.p.A.", Italian National Project under Industrial development programmes for strategic production chains – INVITALIA S.p.A. - Decrees of the Italian Ministry of Economic Development (MISE), DM 09 December 2024 – DM 11 may 2023 Title II art. 4 and the art. 6 with art.4 of DD 18 july 2023.
- EU Project EDIH4Marche under the Call DIGITAL-2021-EDIH-01 with proposal number 210834771.
- ELECTROSPINDLE 4.0 Project from the Italian Ministry of Economic Development (MiSE) in the national plan for research and innovation 2018 titled Smart Factory (Fabbrica Intelligente), Project no. F/160038/01-04/X41.
- RP Marche - European Social Fund Plus (ESF+) 2021/2027.
- European Union – Next Generation Eu - under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 1, Investment 4.1 - Decree No. 118 (2nd March 2023) of Italian Ministry of University and Research - Concession Decree No. 2333 (22nd December 2023) of the Italian Ministry of University and Research, Project code D93C23000450005, within the Italian National Program PhD Programme in Autonomous Systems (DAuSy).


## I. Introduction

The accelerated technological progress of recent decades has profoundly transformed consumption habits globally. However, in the current business model, this evolution is environmentally and economically unsustainable [1]. The textile industry has a strong environmental impact, characterised by a low fibre recycling rate, intensive use of non-renewable resources and polluting processes. The rise of fast fashion, with products destined for short life, exacerbates the problem, making the development of sustainable solutions to reduce the consumption of virgin resources urgent [2]. Today, many companies integrate recycling into their production processes for a more sustainable use of natural resources. However, the complexity and variability of products limit their effectiveness. To address this challenge, industries are adopting the circular economy, implementing take-back programmes and remanufacturing processes [1]. The former enable the recovery of components from end-of-life products, while the latter promote non-destructive disassembly, optimising the reuse of resources [1]. One of the main obstacles to the efficient recycling of garments is the presence of heterogeneous components, such as dress zips, buttons and rivets, which have to be separated to ensure the purity of the recycled materials. Traditionally, this disassembly process is manual, laborious and not very scalable, representing a bottleneck for the adoption of circular economy practices in the textile sector. The increasing focus on sustainability in the textile sector has highlighted the need to automate the disassembly processes of garments, in particular to facilitate material recycling. Dress zips, often made of metal or plastic, represent a significant obstacle in the automated separation of textile components, requiring advanced robotic solutions for their efficient disassembly. In recent years, automated disassembly has become increasingly important as robots have demonstrated the ability to perform tasks previously reserved for human intervention, such as

collaborative tasks [3], [4], manual tasks [5], or co-transport [6], [7]. Of particular interest is the disassembly of flexible materials, the robotic handling of which is a complex task. Here, the planning of optimal trajectories is crucial to balance speed and vibration containment in industrial systems. Control issues can be addressed in different ways. Among them, the trapezoidal velocity profile is often used to optimise performance within speed and acceleration limits, but it can generate jerk peaks that trigger mechanical resonances. To reduce such phenomena, input shaping can be applied, which modulates the reference signal through specific pulses. To this regard, Fang et al. [8] proposed a trajectory planning based on predefined segments, which predicts the configuration of the various path sections in advance, resulting in an analytical solution up to fifth order. Furthermore, they introduce an adjustment technique based on vibrational modes to model acceleration profiles and reduce energy in resonance frequencies. Another strategy is to use parametric functions, such as polynomial curves or splines to interpolate the trajectory, which aim to generate motion profiles with good continuity. For example, Fang et al. [9] proposed the use of a jerk profile based on a sigmoid trait function, designed to generate smooth and continuously derivable trajectories while respecting the limits on velocity, acceleration and jerk. In contrast, Lambrechts et al. [10] used feedforward control and fourth-order trajectories to ensure that the force generated is adequate for the required acceleration, while respecting the mechanical limits and capabilities of the actuators.

Despite recent advances in the field of industrial automation and human-robot collaboration, the current literature shows a significant lack of studies dedicated to the the robotic disassembly of garments, and in particular of dress zip. The present work intends to fill this gap through the development of an advanced robotic system capable of autonomously identifying and removing dress zips from sweatshirts by integrating machine vision supported by AI tools that exploit the potential of LLMs into robotic disassembly tasks. These offer greater flexibility and scalability, adapting to different types of garments, different production volumes and reducing the training efforts of the AI in industrial application. The main objective is to ensure to the robot a continuous, precise and adaptable movement to the inherent variability of textile garments, reducing vibration and maximising the robot's operational efficiency [11]. The proposed idea would help to reduce operating time and costs in the long run, thus making textile recycling more economically viable. The developed method consists of two steps: first, an AI-driven framework that relies on LLM detects and localises the garment dress zip; then, the optimised cutting path is generated for a manipulator robot equipped with dedicated cutting tools.

The remainder of this paper is organized as follows: Section II presents the methodology needed for all the steps involved in the process, III describes the properties of the experimental setup, and the results obtained are included in Section IV. Finally, Section V summarizes the properties of the proposed method.

## II. Proposed Methodology

The method was inspired by what a human operator do when removing a dress zip from garment. Generally, the operator, equipped with the necessary cutting tool, carried out following steps: positioning the garment on the work surface with the zip fully visible, identify starting cutting points on both sides of the dress zip and removing the dress zip by making two approximate cuts along its sides manually. The cuts were rough and reduced the amount of fabric recoverable from the rest of the garment. Instead, proposed methodology consist to use of industrial robots to reduce, by specific algorithms, operating time, costs and maximising the amount of recoverable fabric. To replicate the actions performed by an operator the approach combine a multistage AI-based architecture for segmentation and object detection with an algorithm based on jerk signals and a sigmoid function to produce the trajectory. The following sections focus on: image acquisition and dress zip detection, transformation from pixel coordinate to Cartesian space, trajectory planning algorithm and related control implementation.

### A. Image Acquisition and Zip Detection

In this section, a hybrid AI-driven framework for the detection and segmentation of zips and their constituent components zip tape, puller, and bottom stop in garment images, is introduced. The method combines image segmentation using deep learning with object detection guided by a language model, creating a clear and accurate workflow ideal for structured inspections in industrial environments. The integration of LLMs enhances the system's ability to interpret complex visual patterns and adapt flexibly to dynamic industrial environments. In detail, in the first step, Segment Anything Model v2 (SAM2) identifies the general shape and location of the object of interest [12]. SAM2 was chosen for its outstanding flexibility and precision in segmentation tasks, particularly in zero-shot scenarios. Unlike conventional AI-based models that require extensive fine-tuning to adapt to new object classes or domains, SAM2 demonstrates a remarkable ability to generalize and segment previously unseen objects without explicit prior exposure. This eliminates the need for task-specific supervision or parameter updates, making it ideal for dynamic, real-world applications where annotated data is scarce or unavailable [13]. In the second step, a multimodal LLM Gemini by Google, analyses this segmented image to detect specific components and provide structured outputs such as bounding boxes or labels [14]. Gemini was chosen due to the advanced problem solving capability, integrated design and manufacturing workflow creation [15]. The system is designed around a modular, multi-stage architecture (See Figure 1) in which each component is responsible for a distinct subtask, ensuring clarity, traceability, and modular expansion. The process is organised as follows:

1) a raw garment image is acquired using a RGB-D stereo camera ($1920 \times 1080$ px), mounted directly on the wrist of a collaborative robot,

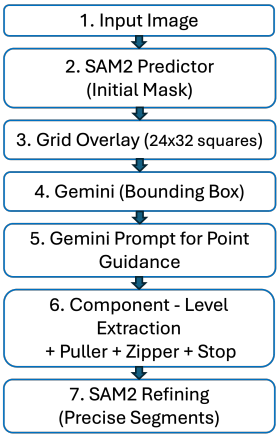



Fig. 1: Pipeline architecture.

2) the image is passed through the Segment Anything Model v2 (SAM2) to produce an initial segmentation mask of prominent structures. This step provides an initial visual isolation of the object without prior training on the specific task.
3) Using a Python script, a coordinate grid is superimposed on the image. It serves as a spatial reference system for the fourth step involving multimodal large language model. The result is a gridded image.
4) The resulting gridded image is used to prompt the LLM, which interprets the visual content and returns structured bounding box predictions for each zip-related component.
5) For each bounding box predicted by the LLM, a second interaction with the model is initiated. The LLM Gemini is prompted again. This time to suggest foreground and background point prompts that lie within or around the detected component in order to improve the quality and accuracy of the masks and to remove artifacts.
6) Using the LLM-generated prompts and the bounding box constraints, SAM2 performs a second, more targeted segmentation for each individual component. This results in high-precision binary masks that better separate the object from the background and eliminate noise or artifacts.
7) The final outputs include color-coded masks, annotated bounding boxes, and a visual comparison between original, detected, and segmented images. These outputs are stored for further analysis or for use in downstream robotic operations.

The effectiveness of the proposed dress zip detection and segmentation approach relies on specific operational assumptions. The dress zip and puller must be metallic to ensure sufficient visual contrast against the fabric. The garment should be laid flat to eliminate folds and shadows that could compromise segmentation quality. Uniform and intense lighting is required to enhance the visibility of structural features. Finally, the dress zip must be centrally positioned and vertically aligned on the inspection plane to optimise the performance of SAM2 in segmenting elongated structures.

*Evaluation Metrics*

*a) Intersection over Union (IoU):* It is used to evaluate the quality of semantic segmentation produced by SAM2, specifically for the zip class [16]. It is computed as the ratio of the intersection to the union of the predicted and ground truth pixels:

$$\text{IoU} = \frac{TP}{TP + FP + FN} \tag{1}$$

Since the evaluation focuses solely on the dress zip class, no averaging across multiple classes is performed.

*b) Bounding Box Detection:* The metric Average Precision (AP) is used to evaluate the object localisation capabilities of Gemini, which provides bounding box predictions in the grid coordinates prior to segmentation. The *Average Precision* (AP) summarises the precision recall curve for the detection of the zip object across multiple IoU thresholds [17]. It is defined as:

$$\text{AP} = \int_0^1 P(R)\, dR \tag{2}$$

where $P(R)$ is the precision as a function of recall $R$. Since only the zip class is considered, no mean across multiple classes is necessary.

## B. Pixel to Robot Base Frame Transformation

Detection process allows to identify a set of points belonging to dress zip in pixel coordinates $(u_i, v_i)$. Manipulator movements refers to robot base frame, thus a coordinate transformation is performed to convert each pair $(u_i, v_i)$ in 3D space coordinates $(x_i, y_i, z_i)$. Interested frames for conversion are three (see Figure 2): camera frame $F_{cam}$, tool frame $F_{tool}$ (at half distance between blades) and robot base frame $F_{base}$. Thus, from pixels to operating space, three transformations are needed: from pixel to camera frame [18], from camera frame to tool frame, and from tool frame to robot base frame [19]. For simplicity of explanation each points detected on dress zip

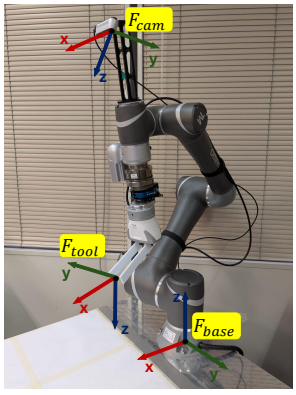



Fig. 2: Positioning and orientation of camera, cutting tool and robot base frames.

will be considered as a generic pixel coordinates $(u, v)$. In the first transformation, the coordinates $(u, v)$ are converted into a 3D vector $p_{cam}$, according to

$$p_{cam} = z_{depth} \cdot K^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}. \quad (3)$$

Where, $z_{depth}$ is the depth for $(u, v)$ measured by RGB-D camera (in metres) and $K$ is intrinsic parameters matrix, with dimension 3x3, containing: focal length, image optical center and screw factor of used camera. To transform the vector $p_{cam}$ into the coordinates of the robot base frame, the last two transformations were computed using homogeneous transformation matrices according to

$$p_{base} = T_{base}^{tool} \times p_{tool} = T_{base}^{tool} \times T_{tool}^{cam} \times \begin{bmatrix} p_{cam} \\ 1 \end{bmatrix}. \quad (4)$$

Where, $T_{base}^{tool}$ is homogeneous transformation matrix from tool frame to base frame (4x4 matrix), $T_{tool}^{cam}$ is homogeneous transformation matrix from camera frame to tool frame (4x4 matrix) and $p_{tool}$ a 3x1 vector containing the desired point expressed in tool frame coordinates. All homogeneous transformation matrices $T_j^i$, from generic frame $F_i$ to generic frame $F_j$, are composed by rotation matrix $R_j$ and translation vector $t_j$ as shown in 5. Where, $R_j$ was obtained by rotation around axes with XYZ sequence $(R_j^z(\vartheta)R_j^y(\varphi)R_j^x(\alpha))$ in fixed frame and the translation vector indicates the position $[x_j^i, y_j^i, z_j^i]$ of $F_j$ respect to $F_i$.

$$T_j^i = \begin{bmatrix} R_j & t_j \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_j^z(\vartheta)R_j^y(\varphi)R_j^x(\alpha) & t_j \\ 0 & 1 \end{bmatrix} \quad (5)$$

The z-component of resulting vector $p_{base}$ was obtained from the difference $(z_{tool} - z_{depth})$, where $z_{tool}$ is the tool position altitude from robot base frame during image acquisition.

### C. Trajectory Planning and Operational Constraints

After capturing a photo with the camera on the robotic arm in the *home position* $p_h$, the dress zip is detected using AI (see Section II-A) and decomposed into coordinates relative to the base frame $F_{base}$ (see Section II-B). The key points are the extremities: the *approach position* $p_a$, where the cut begins (the dress zip point closest to the robot base), and the *final position* $p_f$, where the cut ends (the dress zip point furthest from the robot base). The robotic arm performs two sequential movements: from $p_h$ to $p_a$, then from $p_a$ to $p_f$. During the cutting operation, the robot must execute a smooth motion to achieve clean and continuous cutting edges. Additionally, a fluid trajectory without peaks in force, and thus abrupt accelerations, reduces the risk of tearing the fabric, in contrast to the industrial demand for the shortest possible processing time. Since execution time is crucial in optimal non-linear control [20], this work focusses on online time-optimal trajectory generation. To satisfy these requirements, various time-optimal motion profiles are considered. Initially, trapezoidal profiles for velocity, acceleration, and jerk are analysed [7], [10], followed by trigonometric, sigmoid-based, and elliptic jerk profiles [9], [21]. Fang et al. [8] also compared different S-curve

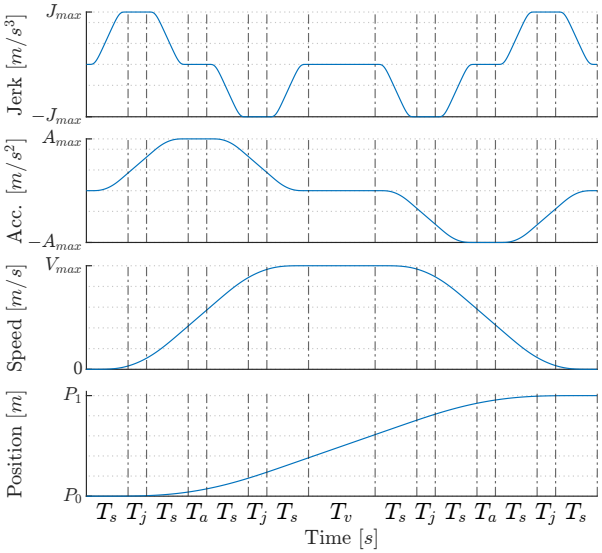



Fig. 3: References generated by the Sigmoid Planner.

profiles, through vibration response analyses and numerical evaluations. These studies showed that smoother high-order derivatives lead to lower measured vibrations. Consequently, the approach using sigmoid-generated jerk signals is adopted. As noted in [9], trajectory planning can be performed either in joint space or in Cartesian space. The Cartesian approach enables a direct comparison between the planned trajectory and the actual zip path, resulting in a more intuitive and effective analysis. Moreover, since the dress zip curve may not be straight, a parametric variable $s$ is introduced, representing the distance along the curve. Thus, the trajectory is computed as a function of $s$ and then converted to Cartesian coordinates. This method is applied to both movements: from $p_h$ to $p_a$, and from $p_a$ to $p_f$. The parametric variable was also used in a previous project [7], but in joint space, requiring more effort to calculate the velocity and acceleration constraints. As time-optimal approaches, speed, acceleration, and jerk must reach their maximum allowable values in the shortest possible time. These limits are selected based on the specific application and operational context. Generally, they are dictated by the hardware specifications; for collaborative robots, they must also comply with ISO safety standards. However, in this specific application where no human operators are present due to the use of a cutting tool, the limits are deliberately reduced to ensure smoother, more controlled cutting.

Once the boundaries are defined, the time intervals $T_s$, $T_j$, $T_a$ and $T_v$ are computed as in [9]. As shown in Figure 3, jerk varies according to a sigmoid function during $T_s$ and remains constant during $T_j$, acceleration is constant during $T_a$, and the speed is constant during $T_v$. These intervals allow the system to move from an initial position $P_0$ with zero speed, acceleration, and jerk, reach the target position $P_1$ while maximizing these quantities within their limits, and conclude again at zero values. Additionally, the end-effector must follow the dress zip not only in position but also with the correct

orientation. To this end, the angles are computed pointwise to align with the cutting line, optimising the operation. While pitch and roll angles remain fixed at zero, for each pair of consecutive points (with indices $k$ and $k+1$) the yaw angle $\psi$ is calculated as:

$$\psi_k = atan2(y_{k+1} - y_k, x_{k+1} - x_k) + \pi. \tag{6}$$

### D. Trajectory Control

A closed-loop control system, based on inverse kinematics, has been designed which allows the optimum approach and cut reference trajectories to be followed. In order to follow the desired trajectory, the developed control calculates the position error $e_P$ (in Cartesian space) and the orientation error $e_O$ (represented by Euler angles) of the end-effector constantly. These errors are used to correct the current position of the robot by (7), which compute, at each step, the joint velocities vector $\dot{q}$,

$$\dot{q} = J_G^{-1}(q) \begin{bmatrix} \dot{p}_d + K_P e_P \\ \dot{\phi}_d + K_O e_O \end{bmatrix}. \tag{7}$$

Where, $J_G^{-1}$ is the inverse of geometrical Jacobian, $\dot{p}_d$ is the desired translation velocity, $\dot{\phi}_d$ is the desired angular velocity and $K_P, K_O$ are defined positive matrices.

## III. Experimental Setup

The approach shown in Section II has been developed and tested in simulation on two computers. First, the AI framework was developed in Python and executed on a MacBook Pro equipped with an Apple M1 Max $3.2\,GHz$ CPU, $32\,GB$ of RAM, 10 core. Then, trajectory generation and control algorithm was developed in *MATLAB®*, using a laptop with 13th generation Intel i7 $2.10\,GHz$ CPU and $32\,GB$ RAM. In simulation a sampling time $t_s = 0.05s$ and a pause time $t_p = 1s$ were used between the end of the approach movement and the beginning of cutting operation. The pause time is assumed to ensure the robot is able to prepare the cutting tool and start the cutting operation. A Collaborative Robot manipulator Omron TM5-900 6DOF fixed on the side of a horizontal work surface was considered. The manipulator was equipped with an Intel RealSense D435i depth camera, precisely through a rigid support on the wrist (link 5). The robot's movement for image acquisition and AI-based processing, as well as the information orchestration between all systems is managed by ROS2 with the architecture described in [22]. The robot is also able to monitor its own failures and anomalies compared to planned behaviour [23], [24]. Furthermore, a cutting tool, with double parallel blades (adjustable in distance), was hooked to the end-effector. This allows more flexibility in the zip formats to be processed due to the adjustable intermediate distance between the blades through the adaptive opening of the terminal fingers. Experimental set-up in virtual simulation and laboratory environment is shown in Figure 4 with end-effector place in the *home position* $p_h = [0.1436, -0.1223, 0.9560]^T$, which enables the camera to take the picture of all the interested

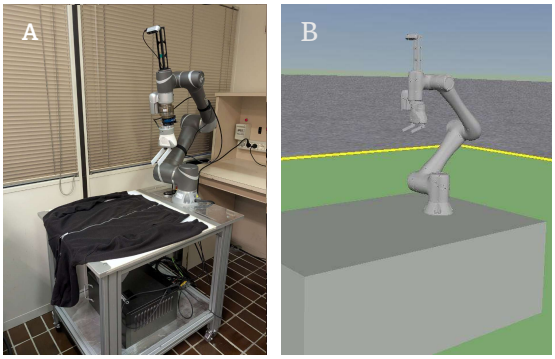



Fig. 4: A) Experimental setup in laboratory. B) Experimental setup in virtual environment.

region. Before the image acquisition, the operator placed the zipped garment with a cut-resistant plate immediately behind the dress zip to avoids tissue dissection of the garment rear side. As shown in the previous Figure 4-A, dress zip must to be set in the center of cutting table. The garment is checked to be laid flat and the surrounding environment is equipped with lights to ensure complete visibility of the dress zip, as discussed in Section II-A. The dress must be placed at 35 cm distance from the robot base and with the zip vertical compared to the limit of positioning, this prevents the robot from kinematic singularity. In addition, the garment must be locked with stops on the cutting table so that the cut does not compromise the fabric.

## IV. Results and Discussion

The experimental evaluation of the proposed pipeline was conducted on an image of a garment containing a visible dress zip structure (Figure 5-A).

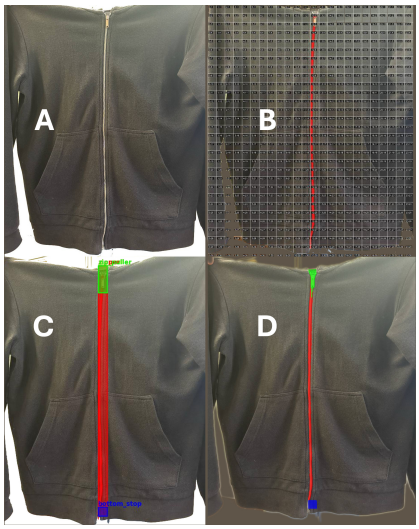



Fig. 5: A) Original image of the sweatshirt with zip. B) Image overlaid with grid used for bounding box localization. C) Annotated result with zip, puller, and bottom stop segmented. D) Annotated high precision binary mask.

The raw input depicts a zipped sweatshirt, serving as a representative example of a real-world industrial use case. After applying the SAM2 model, the extracted segmentation mask was used as a foundation for the subsequent grid generation stage, which overlays a structured $24 \times 32$ coordinate system over the image (Figure 5-B). The annotated output image, as highlighted in Figure 5-C, displays the bounding boxes of the three main components predicted by the Gemini: the full dress zip, the zip puller (slider and pull tab), and the bottom stop. Finally, Figure 5-D, highlights a high precision binary mask that better separates the object from the background and eliminates noise or artifcats. The evaluation was conducted

| Metric | Model | Result |
|---|---|---|
| Intersection over Union | SAM2 | 0.74 |
| Average Precision | Gemini | 0.71 |

TABLE I: Summary of evaluation metrics for segmentation and bounding box detection on the test dataset.

on a dataset consisting of 50 images, each containing a single zip instance. As shown in Table I, for the semantic segmentation task performed by SAM2, the IoU for the zip class achieved a mean value of $0.74$, suggesting moderate segmentation performance with occasional inaccuracies around the edges of the zip, especially in cases with wrinkles or inconsistent lighting conditions. Instead, in terms of Gemini bounding box detection, the AP reached $0.71$, indicating that most zip instances were correctly localised, although some bounding boxes exhibited noticeable shifts or partial overlaps, particularly in images where the zip was not perfectly vertical or the metallic contrast with the garment was reduced.

The extracted segmentation mask of dress zip was used as information to extract points to generate cutting trajectory. Since the extracted mask represents a surface instead of a line, reduction has been adopted. From the upper zip of the dress to the bottom, each 5 pixels were considered pixels in the central position, in order to reduce the width. The coordinates of each pixel were then converted as explained in Section II-B. The Cartesian coordinates of the key points are the following: $p_h$ as defined in Section III, $p_a$ in $[0.3471, -0.0411, 0.4999]^T$ and $p_f$ in $[0.9771, 0.0492, 0.5002]^T$.

Once the dress zip is decomposed in a set of points, transformed in Cartesian coordinate with respect to the base frame $F_{base}$, it is possible to plan the trajectory of the end effector. As specified in Section II-C, the trajectory is planned for the two steps: first from $p_h$ to $p_a$, then after a time $t_p$ from $p_a$ to $p_f$. The interval $t_p$ is needed for the robotic arm to prepare the cutting tools on the sweatshirt, here assumed to be $t_p = 1\,s$.

In this case study, the value of the snap is set to $S_M = 125\,m/s^4$; the jerk is bounded with an absolute amplitude $J_M = 7.5\,m/s^3$, the acceleration with $a_M = 2.87\,m/s^2$ and the speed with $v_M = 0.25\,m/s$. The boundaries of the snap and jerk are the same of the second set in the first numerical example reported in [9], while the acceleration and the speed are reduced for a higher care and control of the

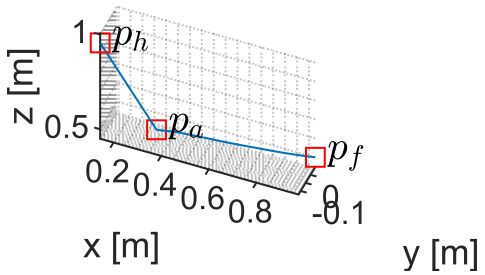


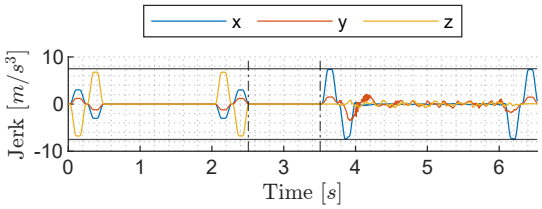

Fig. 6: Reference for the end-effector

cutting operation. Figure 6 shows the path of the end-effector in the three dimensional Cartesian space, linking the three key points.

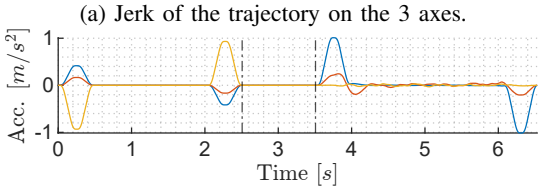
Figure 7, instead, illustrates the path during the time, namely the planned trajectory, decomposed along each axis. The figure highlights each velocity, acceleration, and jerk profiles, split into three parts by dashed lines: the first movement ($0\,s$ - $2.51\,s$), the preparation of the cutting tool ($2.51\,s$ - $3.51\,s$) and the second movement ($3.51\,s$ - $6.5394\,s$). The planned trajectory allows reaching all the key points without overcoming any of the signals boundaries: due to the high snap limit, the jerk consistently reaches its maximum value $J_M$; the acceleration profiles do not attain the upper bound $a_M$, as they are constrained by the relatively low velocity limit $v_M$. Moreover, the position and speed signals are smooth, the speed presents small oscillations, due to the curves of the dress zip. These speed perturbations are the causes of the acceleration oscillating signal; the most clear is with the $y$ signal in Figure 7b, whose perturbations are, however, bounded in the small range $[-0.2\,m/s^2, 0.25\,m/s^2]$. Similarly, the jerk is affected showing perturbations, also in this case the most clear is with the $y$ signal in Figure 7a, and they results bounded in $[-3.6\,m/s^3, 2.2\,m/s^3]$. This analysis lets us know that the robot will be affected by oscillations but with little impact.

The planned trajectory is used as reference for controlling the robotic arm. The simulation is carried out with the trajectory control introduced in Section II-D and the sampling time $t_s$ defined in Section III. During the simulation the data corresponding to the joint variables are acquired and are presented in Figure 8. The evolution of the joint variables throughout the simulation is smooth, with the sole exception occurring at the beginning of the cutting operation, after $t = 3.51\,s$, due to the greater variations along the trajectory (see Figure 8d). The joint speed, acceleration and jerk (see Figures 8c, 8b and 8a, respectively) confirm that the most challenging segment of the trajectory is at the beginning of the cutting operation. While these signals exhibit several consecutive peaks during this phase, they remain confined within narrow acceptable ranges.

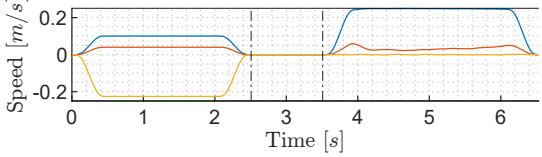
As shown in Figure 9, the robotic arm control system allows to follow the reference trajectory accurately. The end-effector exhibits minimal misalignment: the maximum positional deviation is $4.31\,mm$ along the x-axes and $4.14\,mm$ along the z-axes, while the maximum rotational deviation is



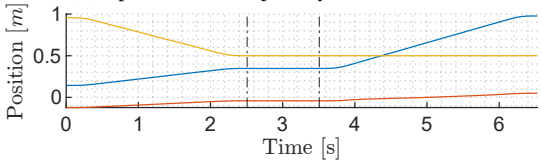

(a) Jerk of the trajectory on the 3 axes.




(b) Acceleration of the trajectory on the 3 axes.

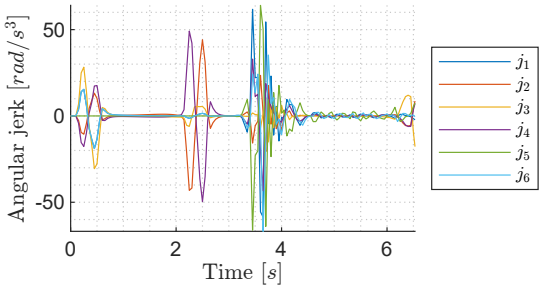



(c) Speed of the trajectory on the 3 axes.

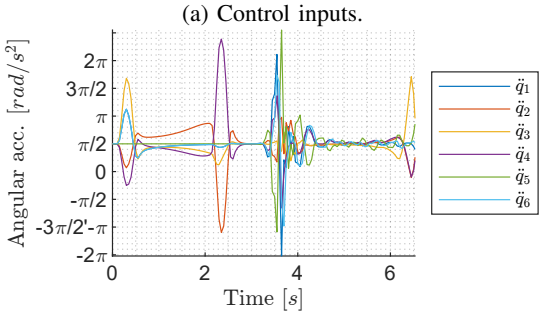



(d) Position of the trajectory on the 3 axes.

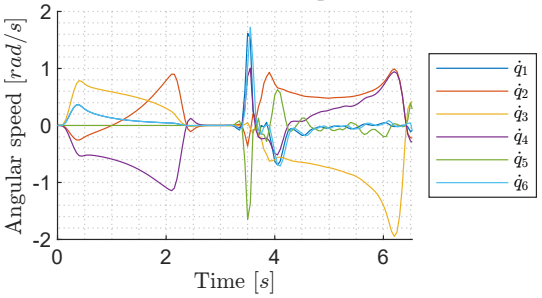
Fig. 7: The trajectory on the 3 axes.



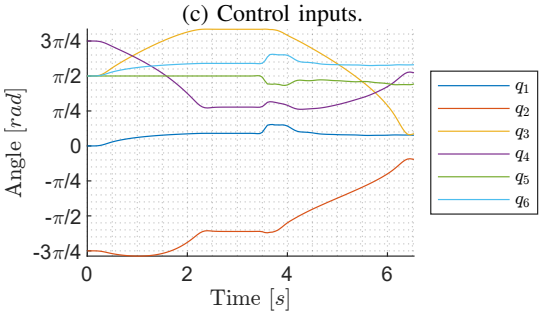

(a) Control inputs.




(b) Control inputs.




(c) Control inputs.




(d) Joint state.

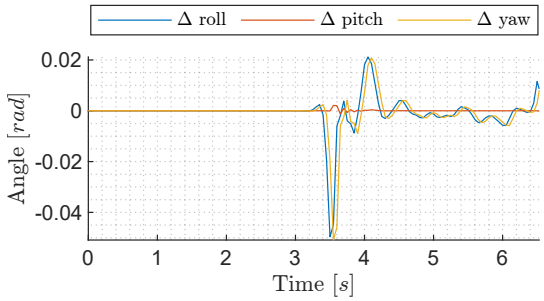
Fig. 8: Analysis of the planned path during cutting operation.

$0.05\,rad \approx 2.86°$, occurring in the roll and pitch direction. Focusing in Figure 9b, it should be noted that the trajectory error in x-axes cutting direction is not due to imprecision in execution. But in actuality, it is due to the time delay in pursuing the desired trajectory. This is because the objective explained algorithm in Section II-C is to minimise trajectory execution delay time.

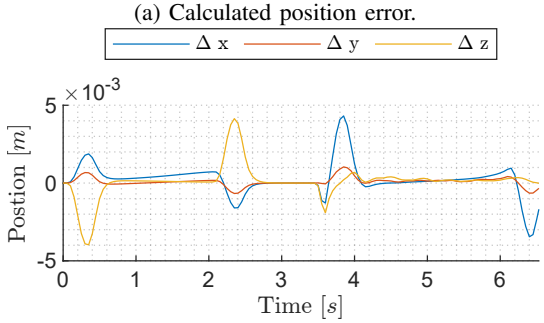
Indeed, the executed path has a minimal error. Figure 10 illustrates the executed trajectory in Cartesian space alongside the sample points of the dress zip, demonstrating that the cut remains centered on the set of reference points. This condition ensures that the recovered material remains uncontaminated and free from unwanted components that could compromise its quality.

## V. Conclusions and Future Developments

In the coming decades, the transition of the global textile industry towards a circular paradigm will be accelerated by regulatory pressures economic and environmental sustainability requirements, and growing consumer awareness. Despite this, the textile industry currently shows limited preparedness for this change due to a lack of expertise in recycling-orientated design and the absence of established technologies for the efficient use of textile waste. In this context, automated disassembly is crucial for the integration of circular economy principles, but it brings challenges due to the handling of heterogeneous components with high flexibility. The presented work proposes a solution for the automatic removal of clothing zips, using an LLM-based approach as an alternative to classical AI. Identifying the points in the clothes zips from RGB-D images, they are converted to three-dimensional coordinates to generate the cutting path followed by a collaborative robot in the removal task. A trajectory is planned which allows the robotic arm to move from its initial position to precisely follow the profile of the dress zip, ensuring an optimal cutting operation. Although the tracking error is small, oscillations in the joints are observed, which remain within tolerable




(a) Calculated position error.




(b) Calculated angular error.

Fig. 9: Analysis of the planned path during approaching and cutting operation.

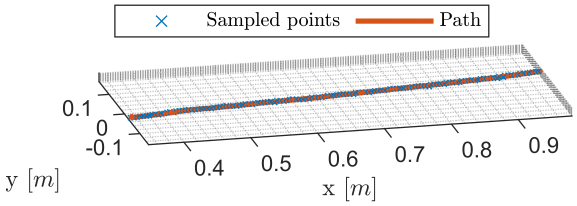



Fig. 10: The sampled points of the dress zip on the executed path.

limits. The experimental results demonstrate the validity of the method, showing that the system is able to correctly execute the path along the identified dress zip.

This study is an important step towards the sustainable automation of textile recycling, supporting the development of a more circular and efficient economy. Future work will focus on optimising trajectory planning and low-level control to further reduce oscillations while maintaining high cutting quality. It will also be necessary to reduce the constraints imposed, to make the system more flexible and adaptable to different types of garment.

## References


[1] S. Hjorth and D. Chrysostomou, "Human–robot collaboration in industrial environments: A literature review on non-destructive disassembly," *Robotics and Computer-Integrated Manufacturing*, vol. 73, 2022.

[2] P. Harmsen, M. Scheffer, and H. Bos, "Textiles for circular fashion: The logic behind recycling options," *Sustainability*, vol. 13, 2021.

[3] L. Wang, R. Gao, J. Váncza, J. Krüger, X. V. Wang, S. Makris, and G. Chryssolouris, "Symbiotic human-robot collaborative assembly," *CIRP annals*, vol. 68, no. 2, pp. 701–726, 2019.

[4] Y. Cheng, L. Sun, C. Liu, and M. Tomizuka, "Towards efficient human-robot collaboration with robust plan recognition and trajectory prediction," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2602–2609, 2020.

[5] R. Kermenov, S. Foix, J. Borràs, V. Castorani, S. Longhi, and A. Bonci, "Automating the hand layup process: On the removal of protective films with collaborative robots," *Robotics and Computer-Integrated Manufacturing*, vol. 93, p. 102899, 2025.

[6] D. Sirintuna, A. Giammarino, and A. Ajoudani, "An object deformation-agnostic framework for human–robot collaborative transportation," *IEEE Transactions on Automation Science and Engineering*, vol. 21, no. 2, pp. 1986–1999, 2023.

[7] R. Kermenov, A. Di Biase, I. Pellicani, S. Longhi, and A. Bonci, "Near time-optimal trajectories with iso standard constraints for human–robot collaboration in fabric co-transportation," *Robotics*, vol. 14, no. 2, 2025.

[8] Y. Fang, C. Gu, Y. Zhao, W. Wang, and X. Guan, "Smooth trajectory generation for industrial machines and robots based on high-order S-curve profiles," *Mechanism and Machine Theory*, vol. 201, p. 105747, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0094114X24001745

[9] Y. Fang, J. Hu, W. Liu, Q. Shao, J. Qi, and Y. Peng, "Smooth and time-optimal S-curve trajectory planning for automated robots and machines," *Mechanism and Machine Theory*, vol. 137, pp. 127–153, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0094114X18317890

[10] P. Lambrechts, M. Boerlage, and M. Steinbuch, "Trajectory planning and feedforward design for electromechanical motion systems," *Control Engineering Practice*, vol. 13, no. 2, pp. 145–157, 2005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0967066104000462

[11] B. Tapia Sal Paz, G. Sorrosal, A. Mancisidor, C. Calleja, and I. Cabanes, "Reinforcement Learning-Based Control for Robotic Flexible Element Disassembly," *Mathematics*, vol. 13, no. 7, 2025. [Online]. Available: https://www.mdpi.com/2227-7390/13/7/1120

[12] Meta AI, "Introducing Meta Segment Anything Model 2 (SAM 2)," 2025. [Online]. Available: https://ai.meta.com/sam2/

[13] J. Reddy, H. Niu, J. L. Scott, M. Bhandari, J. A. Landivar, C. W. Bednarz, and N. Duffield, "Cotton Yield Prediction via UAV-Based Cotton Boll Image Segmentation Using YOLO Model and Segment Anything Model (SAM)," *Remote Sensing*, vol. 16, 12 2024.

[14] Google, "Google AI for Developers," 2025. [Online]. Available: https://ai.google.dev/

[15] C. A. Escobar, J. A. Cantoral-Ceballos´, M. Tworek, K. Yang, and R. Morales-Menendez, "Generative Ai for Manufacturing Innovation: A Comprehensive Review with Case Studies," 2025.

[16] Y.-J. Cho, "Weighted Intersection over Union (wIoU) for evaluating image segmentation," *Pattern Recognition Letters*, vol. 185, pp. 101–107, 2024.

[17] R. Padilla, S. L. Netto, and E. A. B. da Silva, "A Survey on Performance Metrics for Object-Detection Algorithms," in *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 2020, pp. 237–242.

[18] R. Hartley and A. Zisserman, "Camera Models," in *Multiple View Geometry in Computer Vision*, Cambridge, 2004, pp. 153–177.

[19] B. Siciliano, L. Sciavicco, L. Villani, and G. Oriolo, "Kinematics," in *Robotics: Modelling, Planning and Control*. Springer Publishing Company, Incorporated, 2010.

[20] A. Bonci, S. Longhi, G. Nabissi, and G. A. Scala, "Execution time of optimal controls in hard real time, a minimal execution time solution for nonlinear sdre," *IEEE Access*, vol. 8, pp. 158 008–158 025, 2020.

[21] L. Bruzzone and D. Stretti, "Application of Elliptic Jerk Motion Profile to Cartesian Space Position Control of a Serial Robot," *Applied Sciences*, vol. 13, no. 9, 2023. [Online]. Available: https://www.mdpi.com/2076-3417/13/9/5601

[22] A. Bonci, F. Gaudeni, M. C. Giannini, and S. Longhi, "Robot operating system 2 (ros2)-based frameworks for increasing robot autonomy: A survey," *Applied Sciences*, vol. 13, no. 23, 2023.

[23] R. Kermenov, G. Nabissi, S. Longhi, and A. Bonci, "Anomaly detection and concept drift adaptation for dynamic systems: A general method with practical implementation using an industrial collaborative robot," *Sensors*, vol. 23, no. 6, 2023.

[24] G. Nabissi, S. Longhi, and A. Bonci, "Ros-based condition monitoring architecture enabling automatic faults detection in industrial collaborative robots," *Applied Sciences*, vol. 13, no. 1, 2023.