Vision-based human–robot collaboration for wire harness assembly in automotive manufacturing

Citation for the original published paper (version of record):

Wang, H., Salunkhe, O., Hartmann, A. et al (2027). Vision-based human–robot collaboration for wire harness assembly in automotive manufacturing. Robotics and Computer-Integrated Manufacturing, 103. http://dx.doi.org/10.1016/j.rcim.2026.103390

(article starts on next page)

Robotics and Computer–Integrated Manufacturing 103 (2027) 103390

Contents lists available at ScienceDirect

Robotics and Computer-Integrated Manufacturing

journal homepage: www.elsevier.com/locate/rcim

Vision-based human–robot collaboration for wire harness assembly in automotive manufacturing

Hao Wang [a,*], Omkar Salunkhe [a], Annalena Hartmann [b], Sven Ekered [a], Patrick Bründl [b], Jörg Franke [b], Johan Stahre [a], Björn Johansson [a]

[a] Department of Mechanical Engineering, Chalmers University of Technology, Hörsalsvägen 7A, Gothenburg, 412 96, Sweden
[b] Institute for Factory Automation and Production Systems (FAPS), Friedrich-Alexander-Universität Erlangen-Nürnberg, Fürther Straße 246b, Nuremberg, 90429, Germany

ARTICLE INFO

Keywords:
Human–robot collaboration
Computer vision
Assembly automation
Collaborative assembly
Wire harness assembly
Deformable linear object

ABSTRACT

Wire harness assembly exemplifies a non-rigid object assembly task that remains challenging to automate due to object deformability, occlusions, high variability, and workspace constraints. This study introduces a vision-based human–robot collaboration (HRC) framework developed for wire harness assembly in automotive final assembly processes. In this system, the robot performs repetitive, force-intensive tasks, while the human operator handles operations that require dexterity and adaptive decision-making. The proposed approach combines marker-based pose estimation of wire harness components with a hand-triggered control scheme. Detected hand landmarks define a region of interest, enabling intentional, context-aware robot activation. The HRC framework is validated by performing wire harness installation on the vehicle chassis. A two-phase within-subjects experimental study compares manual installation with HRC-assisted installation in both laboratory and industrially relevant environments, corresponding to Technology Readiness Level 4 to 6. The results indicate that robotic assistance significantly reduces localized physical discomfort and physical demand while maintaining a high success rate in wire harness installation. However, overall workload does not differ significantly between HRC and manual conditions. In the HRC condition, both mental demand and average assembly time increase significantly compared to manual assembly. Cycle-time analysis reveals that the robot execution phase accounts for the largest proportion of total time in the collaborative workflow. These findings suggest that vision-based HRC can provide targeted ergonomic benefits for non-rigid object assembly, while also introducing cognitive and temporal trade-offs. Therefore, real-world deployment requires further improvement in interaction fluency and throughput. The proposed method and code are available at https://github.com/HWANG7308/ClampTracking.

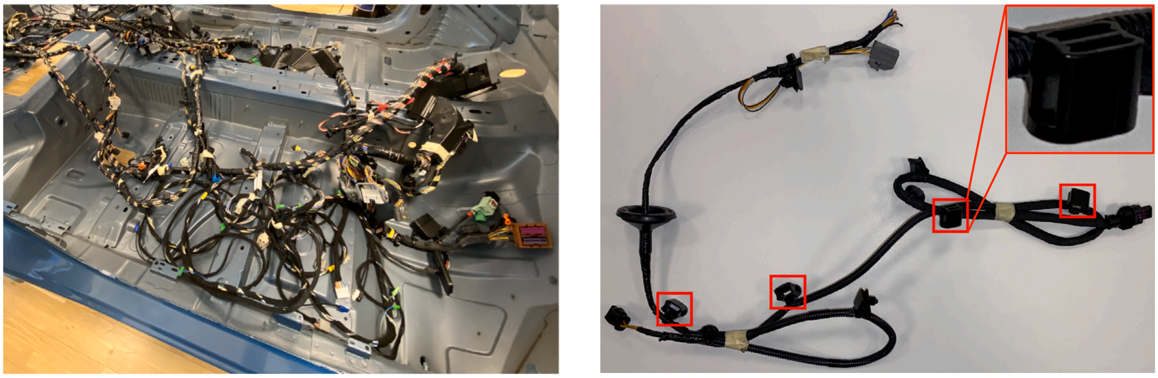
1. Introduction

Factory automation has delivered substantial gains in manufacturing quality and productivity [1]. However, assembly operations remain significantly less automated than other manufacturing processes [2], primarily because they involve complex motion sequences and are strongly influenced by part geometry, tolerances, and contact interactions [3]. These challenges are particularly significant for non-rigid components, as their geometry can change substantially under gravity or external contact forces, complicating perception, modeling, and control [4]. Given these challenges, conventional robot-centric automation frequently proves inadequate [5]. This challenge has motivated the development of hybrid approaches that combine the physical strength and repeatability of machines with human adaptability to achieve robust, efficient assembly performance [6].

Consequently, human–robot collaboration (HRC) is increasingly recognized as a practical approach for industrial automation solutions, especially in high-mix environments with variability and frequent changeovers, which constrain the effectiveness of rigid automation [7]. The advancement of collaborative robots (cobots), their compact form factor, and integrated safety functions further promote their deployment alongside operators in shared workspaces [8]. To fully leverage these systems, effective and robust sensing capabilities must be

* Corresponding author.
E-mail addresses: haowang@chalmers.se (H. Wang), omkar.salunkhe@chalmers.se (O. Salunkhe), annalena.hartmann@faps.fau.de (A. Hartmann), sven.ekered@chalmers.se (S. Ekered), patrick.bruendl@faps.fau.de (P. Bründl), joerg.franke@faps.fau.de (J. Franke), johan.stahre@chalmers.se (J. Stahre), bjorn.johansson@chalmers.se (B. Johansson).

https://doi.org/10.1016/j.rcim.2026.103390
Received 22 January 2026; Received in revised form 27 May 2026; Accepted 13 July 2026
Available online 23 July 2026



(a) (b)

Fig. 1. (a) Example bundle of wire harnesses designed for installation on a car chassis. (b) Section of automotive wire harnesses, with clamps highlighted by red rectangles.

implemented [9]. In particular, vision-based robotic perception is needed to localize relevant objects, interpret operator actions, adapt robot behavior, and enable safe, efficient HRC [10]. Although vision-based methods for manufacturing automation are widely studied, translating them into integrated HRC systems remains challenging [11]. This is especially true for tasks involving deformable or highly variable components, where perception systems must be robust, context-aware, and easy to deploy on real assembly lines [12]. These persistent limitations highlight the need for task-grounded HRC studies that integrate practical perception and interaction design, as well as experimental validation in industrially relevant settings.

In response, this study examines the impact of a vision-based HRC system on the joint performance of a human–robot team during a non-rigid object assembly task. The wire harness assembly process in automotive final assembly is selected as an industrially significant and representative case. As shown in Fig. 1(a), wire harnesses are widely used in products with electrical systems, including automobiles, aircraft, and consumer electronics [13]. From a robotics perspective, wire harnesses can be primarily modeled as deformable linear objects (DLOs) [14]. Their assembly remains predominantly manual due to intrinsic challenges in manipulating DLOs and extrinsic production requirements [5]. Unpredictable shape changes and highly variable configurations increase the difficulty of perception, modeling, and control for conventional robot systems [4]. Additionally, the production context requires flexible task execution rather than tightly constrained fixtures or extensive reprogramming [15].

Previous studies have reported robotic wire harness assembly based on marker-assisted perception [16], model-based localization [17], and model-free recognition [18]. However, these approaches have primarily focused on fully automated or narrowly scoped subtasks under controlled conditions. In contrast, practical deployment in automotive final assembly requires collaborative solutions that accommodate variability, constrained access, and operator-driven task progression. To address this gap, the present study develops a vision-based HRC framework for wire harness installation in which the robot performs the force-intensive pressing action, while the operator performs dexterous handling, positioning, and task sequencing. The proposed framework integrates ArUco marker-based pose estimation [19–21] with an operator-guided hand-triggered interaction mechanism and region of interest (ROI) generation from hand landmarks. Rather than treating these elements as isolated technical components, the framework combines them into a context-aware collaborative workflow. Robot execution is enabled only when the component is deliberately presented in a task-relevant configuration. This design facilitates supportive collaboration within a shared workspace [22] and reduces reliance on rigid fixtures, highly constrained presentation, or extensive reprogramming.

The proposed framework is validated on the wire harness installation task through a within-subjects experimental study. The wire harness assembly process typically involves unpacking the wire harness bundle, positioning it within the vehicle, installing the harness on the chassis by pressing clamps onto welded studs (see Fig. 1(b)), plugging connectors, and inspecting assembly quality [12]. Installing wire harnesses on the chassis represents one of the most physically demanding subtasks in wire harness assembly [23]. This repetitive task frequently requires high localized hand forces and awkward body gestures during clamp insertion, thereby increasing the risk of musculoskeletal disorders (MSDs) [24] and contributing to occupational injuries [25]. In addition, clamp insertion typically demands advanced skills in recognizing, grasping, and manipulating clamps, which are often small, lack texture, and possess intricate structures [26]. These factors make wire harness installation well suited for collaborative automation. In this task, the robot can provide force and repeatability during clamp insertion, while the operator performs dexterous handling and makes contextual decisions about clamp selection, positioning, and sequencing. Fig. 4 presents the overall sequence of the collaborative clamp-insertion process. Subsequent sections provide a detailed explanation of each step.

Experiments were conducted sequentially in both laboratory and industrially relevant environments to correspond with Technology Readiness Level (TRL) stages 4 to 6 [27]. The evaluation incorporates both task outcomes and human-factor measures to quantify the impact of vision-based collaboration on team-level performance during wire harness installation. Accordingly, the core contribution of this work is the development and staged evaluation of a practical vision-based HRC framework for wire harness installation. The contribution does not lie in the isolated use of marker-based pose estimation or hand-triggered interaction as standalone technical elements. The contributions of this work are threefold:

- A practical vision-based HRC framework for wire harness installation that assigns force-intensive clamp insertion to the robot while preserving operator-driven positioning, sequencing, and contextual decision-making;
- An operator-guided, context-aware perception and control strategy that integrates ArUco-based six-degrees-of-freedom (6DoF) pose estimation with hand-landmark-based ROI generation and deliberate robot triggering for collaborative clamp insertion;
- A two-phase, within-subjects experimental evaluation conducted in both laboratory and industrially relevant environments, revealing the trade-off between ergonomic benefit, cognitive demand, throughput, and task effectiveness in a representative non-rigid object assembly task.

2. Related work

Although researchers have investigated automation in wire harness assembly operations, the majority of these processes continue to depend heavily on manual labor [28]. Manual assembly tasks involve repetitive high-force actions, extended reaching, awkward postures, and the handling of heavy wire harness bundles. The combination of variable components, complex geometries, and ergonomic challenges has driven research into collaborative robotic automation strategies aimed at enhancing productivity and operator well-being [6].

2.1. Theoretical categorizations of wire harnesses

From a robotics perspective, wire harnesses belong to the broader class of deformable objects (DOs), which can be categorized as uni-, bi-, or tri-parametric depending on their dominant intrinsic dimensionality [14,29]. Researchers typically model wire harnesses as DLOs [30, 31] or deformable one-dimensional objects (DOOs) [32,33], where one dimension (length) is significantly larger than the other two. Consequently, researchers often study wire harness assembly as a representative instance of robotic manipulation and automated assembly of DLOs [5].

In practical settings, wire harness assembly is more complex than canonical single-DLO manipulation. Wire harnesses exhibit tree-like, branched structures and heterogeneous components [34,35]. These qualities have led to more specific modeling abstractions, such as branched deformable linear objects (BDLOs) [36], DLO networks (DLONs) [37], or deformable multi-linear objects (DMLOs) [38]. Additionally, wire harnesses incorporate rigid elements, including connectors, clips, and clamps, which are attached to deformable cables. This mixed rigidity motivates the category of semi-deformable linear objects (SDLOs) [39].

2.2. Wire harness assembly automation

Early studies explored fully automated robotic systems that manipulate both deformable cables and rigid harness components [6]. Most of the demonstrations occurred in controlled laboratory settings where these setups used stereo and monocular vision [16], fiducial markers on clamps [40,41] or connectors [42,43], and computer-aided-design (CAD) model-based, two-step localization of wire harness bundle and its components [17]. More recently, researchers have explored deep learning-based recognition of wire harness components [44,45] and structural topology [18]. Other contributions address collision-free installation planning based on CAD models [46], hybrid vision-force connector mating [47], sound-based mating detection [48], marker-assisted clamp insertion [40], multifunctional end-effector design for cable manipulation [49], and dual-arm grasping or untangling methods for wire harness branches [50]. While these studies advance robotic technical capabilities, they typically focus on isolated subtasks and rely on assumptions that are difficult to satisfy in real-world production environments, such as stable lighting, unobstructed visibility, simplified backgrounds, and the presence of artificial markers [6].

In production, wire harnesses must be routed through irregular geometry within the vehicle interior. This process involves frequent occlusions, tight clearances, and restricted sensor viewpoints [15]. Cable deformability further complicates perception and deformation estimation, and introduces uncertainty that propagates to grasping and insertion actions [4]. Even advanced vision pipelines struggle with the small sizes and heterogeneous geometries of rigid harness components, as well as the complex, intertwined topology of wire harnesses [12]. Consequently, fully autonomous solutions have yet to be implemented at an industrial scale, and current systems are predominantly limited to laboratory demonstrations or narrowly defined subtasks [5].

These limitations have led to increased interest in HRC as a practical alternative to full automation [51]. Hybrid workflows assign robots to forceful, repetitive, or precise tasks, while humans are responsible for routing, alignment, and dexterous, context-dependent activities [52]. In this division of labor, robots provide strength, repeatability, and precise motion execution, whereas humans contribute dexterity, perceptual robustness, and situational awareness [53]. Collaborative robots enhance operator safety during close-proximity operations and support flexible redeployment across various workstations [8].

For collaborative wire harness assembly, several requirements are repeatedly emphasized in the literature: robust perception of both wire harness state and operator intent [54]; task allocation methods that reflect ergonomic and skill considerations [55]; reliable error detection and recovery mechanisms for mating, routing, and fastening [56]; and compliant end-effectors capable of manipulating rigid and non-rigid components [49]. Furthermore, the confined spaces and limited visibility characteristic of vehicle interiors increase the significance of interaction design and safety constraints when implementing HRC on assembly lines [57].

3. System design and workflow

This section presents the implemented vision-based HRC framework for collaborative wire harness installation. The design does not treat the technical modules as isolated components. Instead, it integrates task allocation, operator-guided perception, and robot execution into a practical collaborative workflow for evaluation in laboratory and industrially relevant environments.

3.1. Prototype robot system

A prototype robotic system was developed to validate essential technologies for vision-based HRC in automotive wire harness assembly, with particular emphasis on the clamp insertion subtask. As shown in Fig. 2, the system consists of a UR3 collaborative robot arm (Universal Robots) positioned in front of the working area. The UR3 achieves pose repeatability of less than 0.1 mm. This precision meets the positioning accuracy requirements for clamp insertion. Robot control is achieved via the Real-Time Data Exchange (RTDE) interface, utilizing the `ur_rtde` Python library [58]. In HRC mode, the robot operates at a reduced speed of 250 mm in accordance with the ISO standard [59].

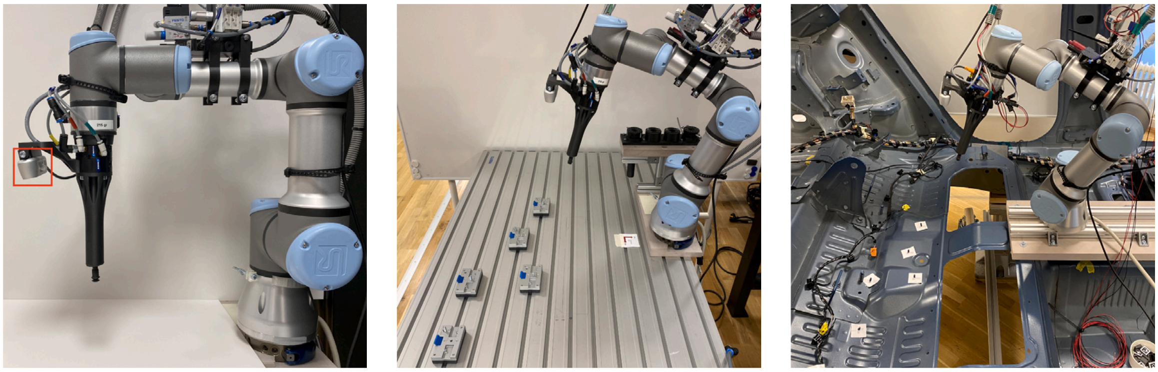
Robot placement is not treated as an optimization problem in this study. A feasible workstation configuration is therefore assumed, allowing the robot to access the operational region and maintain safe collaboration with the operator. The UR3's 500 mm reach does not cover the entire workspace in automotive final assembly, especially for tasks located deep within a vehicle body or spanning the full interior. Addressing these scenarios requires either a robot arm with greater reach or an alternative deployment strategy, such as relocating or repositioning the robot during assembly along the sides of the gantry.

The robot's end effector is fitted with a custom pneumatic tool designed specifically for clamp insertion. The end effector integrates a SCHUNK pneumatic feedthrough and a custom linear actuator. When activated, the pneumatic piston presses the clamp onto the floor-welded stud, enabling the robot to reliably perform the repetitive, force-intensive aspect of the task.

An Intel RealSense D435 RGB-D camera is mounted on the robot flange in an eye-in-hand configuration to support vision-based interaction. The sensor provides aligned RGB and depth images, allowing pixel-wise correspondence between color and depth data. Object detection and pose estimation in this system use fiducial markers attached to the target component, which in this case is the clamp. The RGB stream is used to detect markers and estimate pose, while depth data provides geometric verification and refinement as needed.

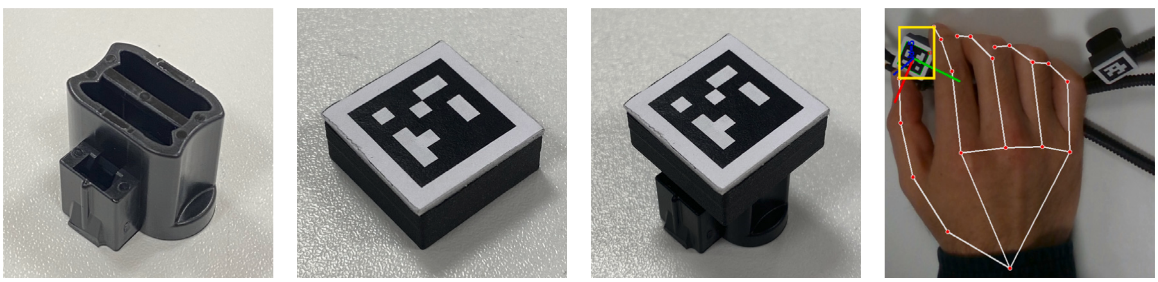
3.2. Clamp cap design

A dedicated clamp cap was developed to ensure reliable clamp detection and pose estimation during collaborative tasks, as shown in Fig. 3. The cap enhances visual observability and improves the repeatability of robotic pressing. It provides a standardized planar surface for fiducial marker-based tracking and remains compatible with clamp insertion. The cap has a cuboidal shape with a square top surface



(a) (b) (c)

Fig. 2. (a) Prototype robotic system equipped with an eye-in-hand camera, highlighted by the red rectangle. (b) System setup on a workbench. (c) System setup within a car body.



(a) (b) (c) (d)

Fig. 3. (a) A wire harness clamp. (b) A designed clamp cap with a marker. (c) A clamp with a cap attached. (d) Hand-triggered marker recognition (ROI is highlighted by the yellow rectangle).

measuring 20 × 20 mm that fully covers the clamp. The cap is 5 mm high, allowing sufficient clearance for marker placement and camera visibility, and preventing interference with insertion or contact with adjacent components. The cap also protects the clamp and surrounding components from damage during automated assembly. The bottom of the cap is specifically designed to match the shape of the clamp. Modifying the cap bottom enables adaptation to other components.

Each cap incorporates a 15 × 15 mm ArUco marker [19–21] centered on its top surface to facilitate robust detection and six-degree-of-freedom (6DoF) pose estimation. Markers are selected from the OpenCV `DICT_5X5_100` dictionary [60], which contains 100 unique identifiers encoded as 5 × 5 binary grids. The fixed marker dimensions and known corner coordinates simplify the perception pipeline and improve repeatability compared with methods that rely only on clamp appearance [16,17,40,41]. Utilizing a standardized marker family also enables the same perception pipeline to be applied to other components without additional feature engineering or retraining of learning-based models.

During operation, the marker is detected in the RGB image by identifying square candidates and decoding the binary pattern to determine the marker ID [19]. The four image-plane corner points are extracted and matched to the corresponding three-dimensional (3D) corner coordinates defined in the marker reference frame.

Using the calibrated camera intrinsics, the marker pose relative to the camera is estimated by solving a perspective-n-point (PnP) problem for the planar target [19]. The pose is obtained by minimizing the reprojection error of the detected marker corners [19]. Because the marker is rigidly attached to the clamp cap, the cap pose is directly determined from this estimate. The clamp pose is then inferred through the known rigid transformation between the cap and the clamp. Hand–eye calibration is subsequently used to transform the estimated pose in the robot base frame, enabling the robot to align its end effector with the clamp and perform the pressurized insertion.

The camera-to-robot transformation was obtained through eye-in-hand hand–eye calibration. A ChArUco calibration board was observed from 20 robot poses. For each pose, the robot TCP pose was recorded through the RTDE interface, while the camera pose relative to the ChArUco board was estimated from detected ChArUco corners using the calibrated RGB camera intrinsics. The hand–eye transformation was then computed using the Tsai method [61] implemented in OpenCV. The resulting transformation was stored and used during operation to convert the estimated marker pose from the camera frame to the robot base frame. The calibration accuracy was evaluated by transforming the estimated ChArUco board poses into the robot base frame and assessing their consistency across validation poses. For each validation pose, the translational deviation was computed as the Euclidean distance between the transformed board position and the mean board position in the robot base frame. The rotational deviation was computed as the angle of the relative rotation between each transformed board orientation and the mean board orientation in the robot base frame, representing the overall 3D orientation deviation. The resulting mean translational deviation was 3.1 mm, and the mean rotational deviation was 1.0°. This accuracy was considered sufficient for the clamp-insertion task because the final insertion relied on the cap-guided pressing operation and the pneumatic end-effector alignment (see Table 1).

3.3. Hand-triggered control

Reliable visual detection and pose estimation require clear target visibility. In wire harness assembly, the inherent flexibility and deformability of cables often lead to occlusions and unpredictable component orientations. As a result, autonomous marker detection is unreliable without external assistance. To address this challenge, a hand-triggered control concept guided by the operator has been introduced. Rather than being treated as an isolated perception module,

Table 1
Key parameters of the vision system and hand–eye calibration.

Purpose	Parameter	Value
Vision system	RGB-D camera	Intel RealSense D435
	Camera configuration	Eye-in-hand, mounted on robot flange
	RGB image resolution	1920 × 1080 px
	Depth image resolution	1280 × 720 px
	Frame rate	30 fps
Clamp cap marker	ArUco [19–21] dictionary	OpenCV `DICT_5X5_100`
	Clamp marker size	15 × 15 mm
	Cap top surface	20 × 20 mm
Finger-defined ROI	Hand landmark detector	Google MediaPipe [62] Hands
	Minimum hand detection confidence	0.5
	ROI definition	Thumb-tip and index-fingertip landmarks
	ROI margin	50 px
Hand–eye calibration	Calibration target	ChArUco board
	ChArUco board layout	5 × 7 squares
	ChArUco square size	30 mm
	ChArUco marker size	15 mm
	Hand–eye calibration method	Tsai method [61] implemented in OpenCV
	Number of calibration poses	20

the hand-triggered mechanism is used here as an enabling interaction strategy that supports deliberate and context-aware collaboration during clamp insertion.

In this approach, the operator holds the clamp-cap combination, aligns it with the designated operation point, and intentionally presents the marker to the camera. During gripping, the marker mounted on the cap is positioned between the thumb and index finger. This action defines the task context by specifying both the component being manipulated and the intended operation location. This approach enables the system to localize and estimate the marker pose exclusively when the marker is intentionally presented, thereby providing an explicit trigger for initiating robotic assistance.

The vision pipeline first performs hand landmark detection on the RGB image using Google MediaPipe [62], which estimates a set of two-dimensional (2D) hand keypoints corresponding to anatomically meaningful landmarks on the detected hand. In the proposed framework, the thumb tip and index fingertip landmarks are used to define the task-relevant interaction region, because these points consistently bound the grasped marker during presentation. A rectangular ROI is generated from the pixel coordinates of these two landmarks by determining their minimum and maximum image coordinates. A fixed margin is then added to the bounding box. This helps keep the full marker inside the search region despite small hand motions, grasp variations, and image noise. Marker detection is subsequently restricted to this ROI instead of being performed over the entire image.

After the ROI is defined, the vision pipeline searches for an ArUco marker [19–21] only within this reduced image region and, when a valid marker is detected, computes its 6DoF pose as described in Section 3.2. Constraining detection to the fingertip-defined ROI reduces the search space and suppresses false positives from irrelevant image regions. It also links marker detection directly to intentional operator actions. Fig. 3(d) shows an example in which only the marker located within the fingertip-defined ROI is detected and used for pose estimation.

This hand-triggered design provides three primary benefits. First, operator-assisted alignment improves marker visibility and reduces perception failures caused by occlusions or unfavorable viewing angles. Second, restricting the search to the ROI improves robustness in cluttered environments by reducing spurious detections outside the task-relevant area. Third, the method enhances safety by permitting robot execution only when the clamp is intentionally grasped and presented, thereby supporting reliable and deliberate collaboration.

3.4. Task allocation and workflow

Task allocation is fundamental to design choices in HRC, especially for complex tasks such as wire harness assembly [10]. A well-defined division of responsibilities minimizes task overlap, reduces coordination overhead, and allows both the human operator and the robot to contribute their respective strengths safely and efficiently. In this study, these design choices collectively define the implemented HRC framework that is subsequently evaluated in terms of ergonomics, cognitive demand, throughput, and task effectiveness.

Building upon the task-allocation framework for collaborative wire harness assembly stations in final assembly [26], the proposed operation first decomposes the installation task into subtasks. These subtasks are then assigned according to their primary cognitive and physical demands. Subtasks that require task comprehension, sequencing, contextual judgment, and dexterous manipulation are assigned to the operator. In contrast, subtasks involving repeatable motion execution and localized force application are assigned to the robot. During the wire harness installation task, the operator identifies the target clamp, retrieves and attaches the cap, positions the clamp-cap combination at the designated insertion point, and maintains the assembly in an appropriate pose for perception. The robot complements the operator's actions by localizing the clamp-cap combination using marker-based perception, estimating its 6DoF pose, assessing reachability, and performing the clamp pressing motion.

This division of labor exemplifies a supportive collaboration pattern within a shared workspace [22]. The operator maintains control over task progression and component presentation, whereas the robot executes the repetitive and force-intensive insertion step. This allocation mitigates the main ergonomic risk in wire harness installation by transferring the repeated localized pressing force from the operator to the robot.

To clarify the system workflow at the module level, the collaborative clamp-insertion cycle is structured as follows:

1. **Task initiation and clamp preparation:** The main control program starts the collaborative cycle for a target clamp. The operator finds the clamp, retrieves a clamp cap from the bin, attaches the cap to the clamp, and places the clamp-cap combination at the desired installation position.
2. **Image acquisition and hand-guided perception trigger:** Once the operator presents the clamp-cap combination, the vision orchestrator requests image data from the camera. The hand detection module processes the RGB image and estimates the operator's hand landmarks.
3. **ROI generation and marker search:** The detected thumb and index fingertip landmarks are used to define a task-relevant ROI. The marker detection module then searches only within this ROI for a valid ArUco marker [19–21] and, upon successful detection, estimates its 6DoF pose in the camera frame.

Algorithm 1 Collaborative clamp-insertion workflow

```
for each target clamp do
    Operator identifies the target clamp
    Operator retrieves a clamp cap and attaches it to the clamp
    Operator places the clamp-cap combination at the target location
    repeat
        Acquire RGB image from camera
        Detect hand landmarks in image
        if hand landmarks detected then
            Define ROI from thumb tip and index fingertip landmarks
            Detect ArUco marker within ROI
        end if
    until valid marker detected within ROI
    Estimate marker 6DoF pose in camera frame
    Infer clamp pose from cap-to-clamp rigid transformation
    Transform clamp pose to robot base frame via hand–eye calibration
    Check clamp reachability
    if clamp is reachable then
        Notify operator that insertion will be executed
        Robot approaches target pose
        Robot aligns tool with clamp
        Robot executes clamp pressing motion
        Robot returns to standby
        Notify operator that pressing is completed
    else
        Notify operator that clamp is not reachable
    end if
    Operator releases the clamp and removes the cap
end for
```

4. **Pose transformation and feasibility check:** The main program receives the hand-landmark and marker-pose information, transforms the marker pose into the robot base frame using the calibrated hand–eye relationship, and evaluates whether the clamp pose is reachable for insertion.
5. **Robot execution:** If the clamp is reachable, the robot controller initiates the insertion sequence. The robot approaches the estimated pose, aligns the pneumatic tool with the clamp, executes the pressing motion, completes the insertion, and then returns to its standby state.
6. **Operator feedback and task continuation:** The system notifies the operator whether the clamp was reachable and, if insertion is completed, informs the operator that the pressing sequence has finished. The operator then releases the clamp, removes the cap, and continues to the next clamp.

Algorithmically, the workflow adheres to a perception–decision–action structure. This formulation clarifies how the proposed HRC framework operationalizes collaborative task execution: perception establishes task context, decision logic verifies execution feasibility, and robot action is triggered only under a deliberate operator-defined state. Specifically, the perception stage consists of synchronized image acquisition, hand landmark detection, ROI definition, and marker pose estimation. The decision stage includes pose transformation and reachability assessment. The action stage consists of robot alignment, clamp pressing, and returning to standby. Robot execution is initiated only after successful detection of both the operator's hand and a valid marker within the fingertip-defined ROI, thereby explicitly linking robot assistance to a deliberate operator action.

Algorithm 1 and Fig. 4 summarize the resulting module-level interactions and execution logic of the proposed collaborative clamp-insertion workflow.

4. Experiments

4.1. Experimental design

A hypothesis-driven experimental design [63] was employed to quantitatively evaluate the ergonomic and performance impacts of the proposed vision-based HRC system for wire harness installation. A two-phase, within-subjects design was implemented to incrementally increase task realism in accordance with the TRL scale [27]. The first phase was conducted in a controlled laboratory environment at TRL 4 (**P1**), followed by a second phase in an industrially relevant environment using a real car body, corresponding to TRL 5 to 6 (**P2**).

All participants completed both phases in a fixed sequence (**P1** → **P2**). This sequence was chosen for two reasons. First, it prevented participants from entering the more constrained car-body environment without prior experience. Second, it followed the TRL scale [27] on validating the system first in a laboratory setting and then in an industrially relevant environment. Within each phase, participants performed wire harness installation under two control conditions: (1) a fully manual condition (**C1**) and (2) a HRC condition (**C2**), in which the robot executed the pressing operation and the participant completed the remaining actions. To minimize order and learning effects within each phase, the sequence of **C1** and **C2** was counterbalanced among participants. Participants were randomly assigned to either the manual-to-HRC sequence (**S1**) or the HRC-to-manual sequence (**S2**).

The laboratory environment setup (Fig. 5(a)) established controlled conditions and approximated floor-weld stud locations and reach characteristics to emulate in-vehicle industrial assembly. The industrially relevant setup (Fig. 5(c)) integrated the robotic system with a real car body, replicating the spatial constraints, limited visibility, and posture requirements characteristic of automotive final assembly. To ensure experimental consistency, the robot was mounted in a fixed configuration. In production settings, alternative mounting or repositioning strategies may be implemented to increase the robot's reachable workspace. Figs. 5(b) and 5(d) illustrate HRC in each scenario.

In each phase and condition, each participant completed five clamp insertions. This resulted in 20 trials per participant (2 phases × 2 conditions × 5 repetitions). The primary independent variable was the control condition (**C1** versus **C2**), evaluated in two experimental phases (**P1** and **P2**).

The following hypotheses guided the experiments:

- **H1**: Localized hand discomfort will be lower in **C2** than in **C1** in either phase.
- **H2**: General physical discomfort will be lower in **C2** than in **C1** in **P2**.
- **H3**: Perceived workload will be lower in **C2** than in **C1** in either phase.
- **H4**: Assembly time will not differ significantly between **C1** and **C2** in either phase.
- **H5**: Success rate will not differ significantly between **C1** and **C2** in either phase.

4.2. Participants

Twenty adults participated in the study (N_p = 20; 11 female, 9 male), with a median age group of 25 to 34 years. The sample size is consistent with previous experimental studies in HRC and human–robot interaction (HRI) that assess human factors and task performance outcomes in controlled and semi-realistic environments [64–67].

Participants were recruited from the local university and surrounding community using convenience sampling [68]. Inclusion criteria required normal or corrected-to-normal vision and sufficient manual dexterity to perform repeated wire harness handling and alignment tasks. To minimize risk and potential confounding in discomfort assessments, individuals with current upper-limb MSDs, acute hand or

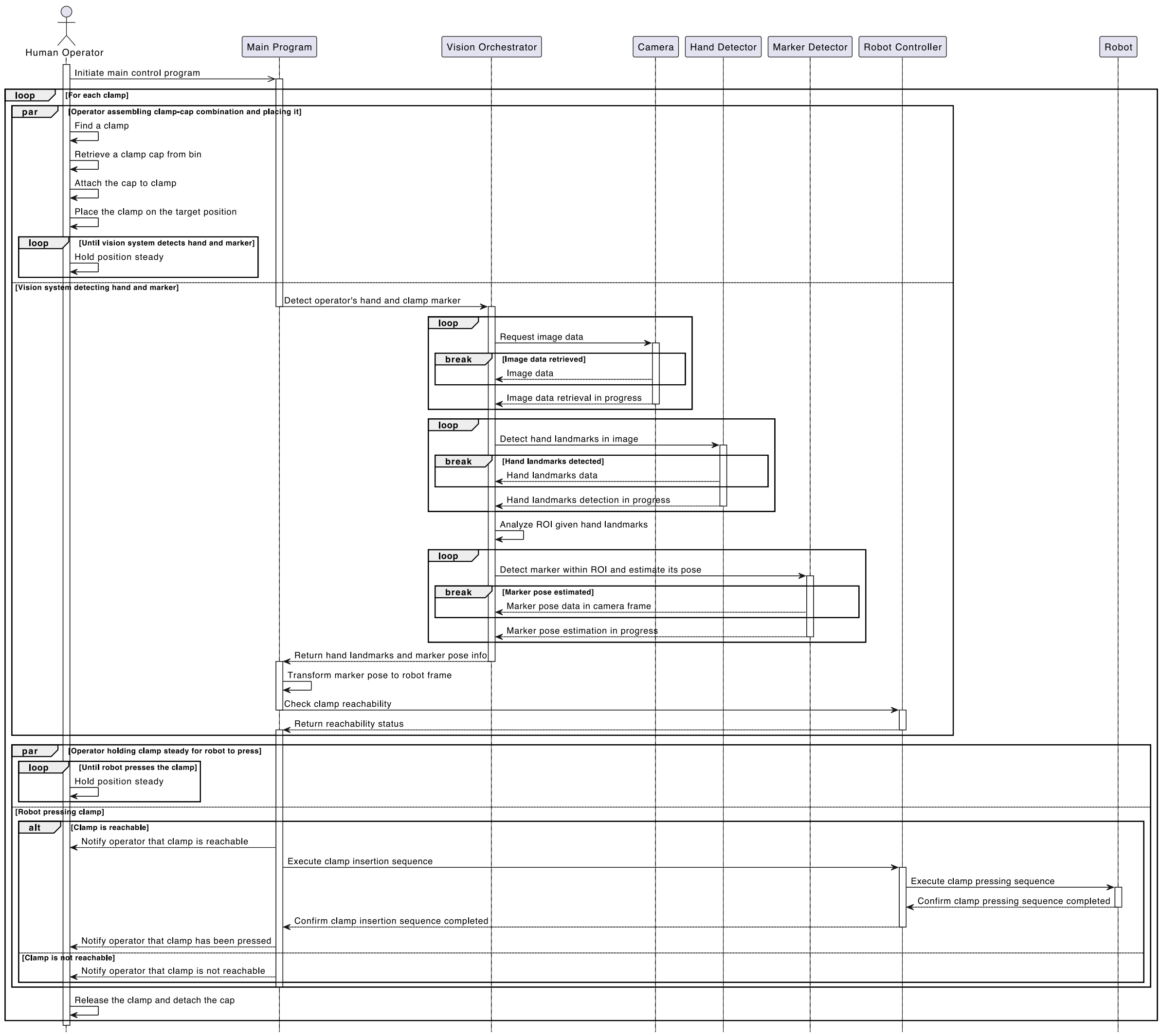
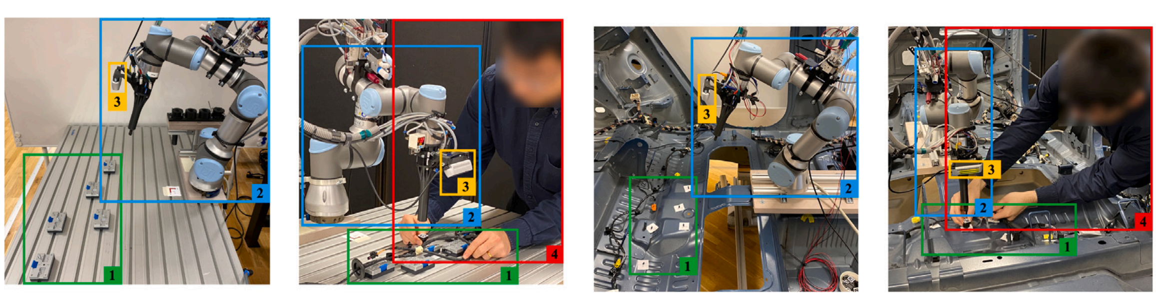


Fig. 4. UML sequence diagram illustrating system module interactions during collaborative wire harness clamp insertion.



(a) (b) (c) (d)

Fig. 5. Experimental setups. (a) Laboratory environment setup. (b) HRC in the laboratory environment. (c) Industrially relevant environment setup. (d) HRC in the industrially relevant environment. Colored rectangles indicate the following regions in each image: 1: shared working region, 2: robot, 3: camera, 4: operator.

finger pain, or neurological or orthopedic conditions affecting grip or dexterity were excluded.

No restrictions were imposed on participants' backgrounds or prior experience with robotics or assembly. This resulted in a participant pool with diverse experience levels, which partially reflects the variability in operator experience encountered at the targeted final assembly station. Fig. 6 presents participant demographics, including gender, age group, and self-reported familiarity with assembly, artificial intelligence, computer vision, robotics, HRC, and wire harness assembly.

4.3. Measurements and scales

The primary independent variable is the control condition, which is defined by whether clamp insertion is performed fully manually (**C1**)

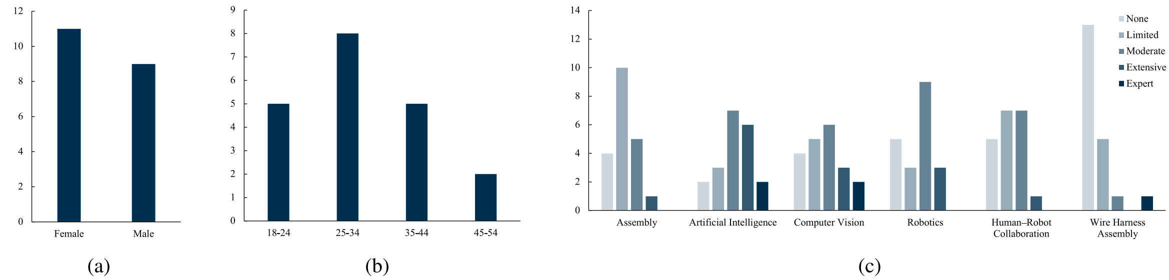


Fig. 6. Participant demographic statistics: (a) gender, (b) age group, and (c) level of expertise.

or collaboratively with robotic assistance (**C2**). Dependent variables include both subjective human-factor measures and objective task-performance metrics, selected to assess ergonomic impact, efficiency, and effectiveness.

1. **Discomfort:** After each experimental block, participants rated their perceived discomfort on a scale from 0 (no discomfort) to 10 (extreme discomfort). Each block consisted of five clamp insertions under one condition within one phase. Two distinct discomfort measures were collected: (i) localized hand discomfort (in both **P1** and **P2**); and (ii) general physical discomfort (collected only in **P2**), which reflects the increased posture and reach demands characteristic of the car-body environment.
2. **Perceived workload:** Subjective workload was assessed after each experimental block using the NASA Task Load Index (NASA-TLX), which evaluates perceived workload across six dimensions: Mental Demand (MD), Physical Demand (PD), Temporal Demand (TD), Performance (OP), Effort (EF), and Frustration (FR) [69]. Participants rated the standard TLX subscales. Overall workload (OW) scores were calculated using both raw (OW_r) and weighted (OW_w) NASA-TLX procedures [70].
3. **Assembly time:** Task efficiency was quantified as the average time required to complete a single clamp insertion within each block. In **C1**, insertion time was measured using a stopwatch. For each clamp insertion, timing began when the participant initiated the insertion attempt and ended when the clamp was fully inserted into the stud. In **C2**, insertion time was also recorded using a stopwatch. Additionally, automated logs captured system-related timing components, including vision processing and robot execution, to enable more detailed analysis of the collaborative cycle. For each clamp insertion, timing began when the participant initiated assembly of the clamp-cap combination and ended when the clamp was pressed by the robot end effector.
4. **Success rate:** Task effectiveness was measured as a binary outcome for each clamp insertion (success or failure). A trial was successful if the clamp was fully seated at the intended location; otherwise, it was failed. The success rate was computed per block as the proportion of successful insertions across the five trials.

4.4. Experiment procedure

Upon arrival, participants received a standardized briefing detailing the study procedures. Before participation, all individuals received both written and verbal explanations of the study objectives, procedures, and potential risks. Written informed consent was obtained from all participants. Participants were informed of their right to withdraw from the study at any time without penalty.

Participants completed a demographic questionnaire and provided a baseline assessment of localized hand discomfort. Safety stop rules were established to safeguard participant well-being. A condition block was terminated immediately if a participant reported excessive discomfort or if visible signs of strain were detected.

Before the experimental trials, participants completed a familiarization session covering both experimental conditions. Participants practiced wire harness installation under the manual condition (**C1**) and the HRC condition (**C2**) until they demonstrated comprehension of the task, the interaction sequence, and robot behavior, and confirmed readiness to proceed. During familiarization, the experimenter reviewed safe interaction practices, including hand placement and stop behavior, and verified that participants could reliably activate the robot using the designated hand-guidance mechanism.

After preparation, each participant completed the two-phase experimental protocol. In **P1**, participants completed two condition blocks (**C1** and **C2**) in a counterbalanced sequence (**S1** or **S2**). After each block, participants completed the NASA-TLX [69,70] workload assessment and reported localized hand discomfort. The same procedure was repeated in **P2** using the car-body setup. In **P2**, participants additionally reported general physical discomfort after each block. Throughout the experiment, the experimenter recorded each participant's assembly time and success rate.

After both phases, each participant was debriefed and asked whether they had experienced any pain or adverse effects during the study. The experimenter also examined the participant for visible signs of irritation or strain. Each participant was thanked for their contribution at the conclusion of the study.

To ensure ethical research, all study procedures complied with institutional ethical guidelines and applicable data protection regulations. No video or photographic material was recorded. All datasets were stored on secure institutional servers in accordance with GDPR and local data protection requirements. To maintain confidentiality, all collected data were anonymized using participant identification codes prior to storage.

5. Results

5.1. Tests of normality

For each dependent variable defined in Section 4.3, distributional assumptions were initially evaluated to identify suitable paired statistical tests for comparing the two control conditions. Given the within-subjects design and a sample size of $N_p = 20$, normality of the paired differences ($\Delta = C2 - C1$) was assessed using the Shapiro–Wilk test. A significance level of $\alpha = .05$ was applied to all statistical tests in this study. Table 2 reports the normality test statistics and corresponding p-values.

Variables with paired differences that did not significantly deviate from normality ($p \geq .05$, indicated as "Not rejected" in the "Normality" column of Table 2) were analyzed using paired-samples t-tests. For variables with paired differences that deviated from normality ($p < .05$, indicated as "Rejected" in the "Normality" column of Table 2), the Wilcoxon signed-rank test was employed as a non-parametric alternative.

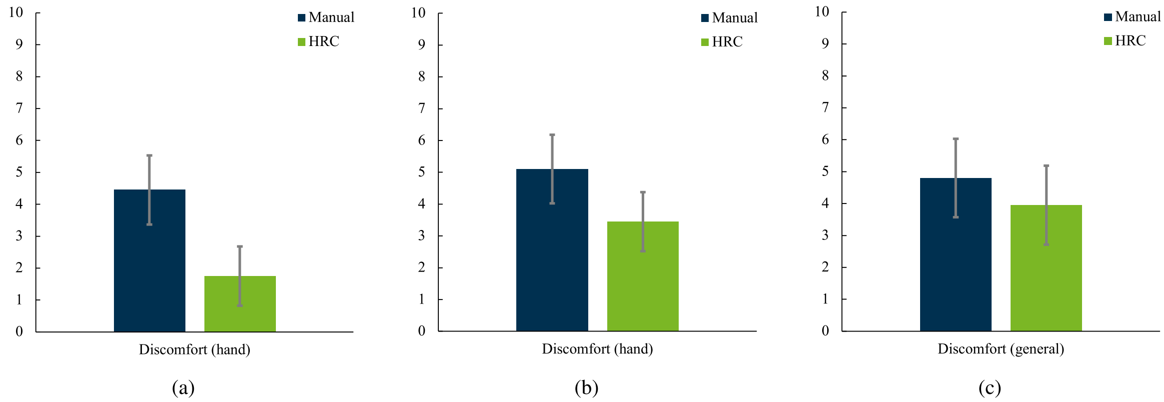


Fig. 7. Discomfort ratings. (a) Localized hand discomfort rating for **P1**. (b) Localized hand discomfort rating for **P2**. (c) General physical discomfort rating for **P2**.

Table 2
Shapiro–Wilk tests of normality for difference scores in **P1** and **P2**.

Variable	P1			P2		
	W	*p*	Normality	*W*	*p*	Normality
Discomfort_{hand}	.963	.607	Not rejected	.961	.561	Not rejected
$\text{Discomfort}_{general}$	–	–	–	.958	.509	Not rejected
NASA-TLX MD	.862	.008	Rejected	.928	.141	Not rejected
NASA-TLX PD	.955	.455	Not rejected	.923	.112	Not rejected
NASA-TLX TD	.935	.194	Not rejected	.895	.033	Rejected
NASA-TLX OP	.933	.174	Not rejected	.872	.013	Rejected
NASA-TLX EF	.949	.351	Not rejected	.943	.278	Not rejected
NASA-TLX FR	.927	.138	Not rejected	.949	.359	Not rejected
NASA-TLX OW_r	.858	.007	Rejected	.912	.069	Not rejected
NASA-TLX OW_w	.851	.006	Rejected	.884	.021	Rejected
Time_{total}	.890	.027	Rejected	.921	.104	Not rejected
Time_{human}	.983	.962	Not rejected	.934	.186	Not rejected
Success rate	.351	<.001	Rejected	.351	<.001	Rejected

Note. The sample size per setting is $N_p = 20$. A value of $p < .05$ denotes a significant departure from normality. Dashes represent variables that were not measured in the respective setting.

5.2. Discomfort rating

Paired-samples t-tests were performed to assess **H1** by comparing localized hand discomfort after manual assembly and the HRC condition in **P1** and **P2**. Figs. 7(a) and (b) present the results, including the upper and lower bounds of the 95% confidence interval (CI) for both experimental phases. In both phases, participants experienced significantly lower localized hand discomfort after HRC assembly compared to the manual condition (see Tables 3 and 5).

A paired-samples t-test was also performed to assess **H2** by examining general physical discomfort in **P2**. Fig. 7(c) presents the results, including the upper and lower bounds of the 95% CI for both experimental phases. Although the mean general physical discomfort was lower in the HRC condition than in the manual condition, this difference was not statistically significant (see Table 5).

5.3. Perceived workload

Perceived workload was evaluated to test **H3** by analyzing the OW and each subscale of the NASA-TLX [69,70] for both **P1** and **P2**. Fig. 8 presents the results, including the upper and lower bounds of the 95% CI for both experimental phases.

In **P1**, nonparametric analyses revealed no significant difference in OW between the manual and HRC conditions; however, MD was significantly higher in the HRC condition (see Table 4). Parametric tests indicated lower PD in the HRC condition, OP favored the manual condition, and no significant differences in TD, EF, and FR (see Table 3).

In **P2**, OW did not differ significantly between the conditions. No significant differences were observed in TD, OP, EF, or FR between the conditions. In contrast, MD was higher, while PD was lower in the HRC condition (see Tables 5 and 6).

5.4. Assembly time

The average cycle time per clamp insertion was analyzed in both **P1** and **P2** to evaluate **H4**. The composition of the cycle time in the HRC condition was also examined to identify the primary contributors to the overall cycle duration. The cycle was divided into three components: (i) human time, which included assembling the clamp-cap combination, positioning, presenting, and triggering; (ii) vision time, which encompassed hand detection, ROI definition, marker detection, and pose estimation by the vision system; and (iii) robot time, which involved the robot's approach, alignment, pressing, and return to standby. Fig. 9 presents the results, including the upper and lower bounds of the 95% CI for both experimental phases.

In **P1**, the Wilcoxon signed-rank test indicated that the average cycle time per insertion was significantly longer in the HRC condition compared to the manual condition (see Table 4). In contrast, the paired-samples t test showed that the average human time per insertion did not differ significantly between conditions (see Table 3).

In **P2**, paired-samples t tests demonstrated that both the average total time per insertion and the average human time per insertion were significantly longer in the HRC condition (see Table 5).

In both **P1** and **P2**, robot execution accounted for the largest proportion of the HRC cycle time, while vision processing contributed only a minor share.

5.5. Success rate

Hypothesis **H5** was evaluated by comparing the success rates between manual and HRC conditions in both experimental phases. Fig. 10 presents the results, including the upper and lower bounds of the 95% CI for both experimental phases. In each phase, Wilcoxon signed-rank tests indicated no statistically significant differences in success rates between the conditions (see Tables 4 and 6).

5.6. Perception robustness

Previous studies have shown that the accuracy and robustness of fiducial-marker-based pose estimation are affected by imaging conditions, marker geometry, viewing angle, marker location in the image, and corner localization quality [71,72]. Experimental evaluations of ArUco and related fiducial markers further indicate that degraded illumination, motion blur, and poor visibility can reduce detection reliability and pose-estimation accuracy [72–74]. Therefore, following

Table 3
Results of paired-samples t tests comparing manual and HRC conditions (**P1**).

Variable	Manual $M(SD)$	HRC $M(SD)$	ΔM [95% CI]	$t(19)$	p	d [95% CI]
Discomfort$_{hand}$	4.450(2.305)	1.750(1.997)	2.700 [1.543, 3.857]	4.883	<.001	1.092 [0.525, 1.640]
NASA-TLX PD	43.000(19.222)	28.000(18.806)	15.000 [4.871, 25.129]	3.099	.006	0.693 [0.196, 1.176]
NASA-TLX TD	37.500(24.682)	38.000(22.850)	−0.500 [−12.502, 11.502]	−0.087	.931	−0.019 [−0.458, 0.419]
NASA-TLX OP	20.000(13.765)	27.500(17.130)	−7.500 [−13.356, −1.644]	−2.680	.015	−0.599 [−1.070, −0.115]
NASA-TLX EF	41.500(23.902)	34.000(20.365)	7.500 [−4.540, 19.540]	1.304	.208	0.292 [−0.160, 0.736]
NASA-TLX FR	31.500(27.582)	30.000(26.358)	1.500 [−10.862, 13.862]	0.254	.802	0.057 [−0.383, 0.495]
Time$_{human}$	4.602(1.831)	5.575(2.405)	−0.973 [−1.971, 0.025]	−2.040	.055	−0.456 [−0.912, 0.011]

Note. The sample size per setting is $N_p = 20$. $\Delta M = M_{\text{manual}} - M_{\text{HRC}}$. All p values are two-tailed. Cohen's d and its 95% CI are reported as provided by SPSS for paired samples, standardized using the standard deviation of the difference scores.

Table 4
Results of Wilcoxon signed-rank tests comparing manual and HRC conditions (**P1**).

Variable	Manual Mdn [IQR]	HRC Mdn [IQR]	Z	p	r
NASA-TLX MD	15.000 [10.000,37.500]	45.000 [30.000,50.000]	−2.993	.003	.669
NASA-TLX OW$_r$	30.833 [20.833,43.333]	35.000 [23.750,40.833]	−0.580	.562	.130
NASA-TLX OW$_w$	31.667 [21.333,46.000]	35.667 [26.000,43.167]	−0.149	.881	.033
Time$_{total}$	4.400 [3.028,6.433]	15.545 [14.255,19.635]	−3.920	<.001	.877
Success rate	1.000 [1.000,1.000]	1.000 [1.000,1.000]	−1.414	.157	.316

Note. The sample size per setting is $N_p = 20$. Analyses were conducted using two-tailed Wilcoxon signed-rank tests. Mdn denotes the median; IQR represents the interquartile range (25th to 75th percentile). Effect size is calculated as $r = |Z|/\sqrt{N}$.

Table 5
Results of paired-samples t tests comparing manual and HRC conditions (**P2**).

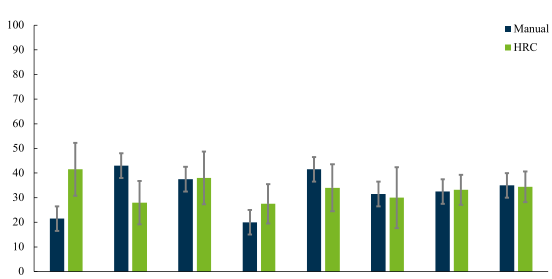
Variable	Manual $M(SD)$	HRC $M(SD)$	ΔM [95% CI]	$t(19)$	p	d [95% CI]
Discomfort$_{hand}$	5.100(2.490)	3.450(2.502)	1.650 [0.307,2.993]	2.571	.019	0.575 [0.094,1.043]
Discomfort$_{general}$	4.800(2.628)	3.950(2.645)	0.850 [−0.616,2.316]	1.213	.240	0.271 [−0.179,0.714]
NASA-TLX MD	28.500(22.308)	41.000(23.373)	−12.500 [−21.206,−3.794]	−3.005	.007	−0.672 [−1.152,−0.178]
NASA-TLX PD	53.500(23.005)	40.500(21.392)	13.000 [2.040,23.960]	2.483	.023	0.555 [0.077,1.021]
NASA-TLX EF	49.000(23.598)	42.000(21.909)	7.000 [−2.969,16.969]	1.470	.158	0.329 [−0.126,0.775]
NASA-TLX FR	33.500(28.149)	33.500(21.095)	0.000 [−8.983,8.983]	0.000	1.000	0.000 [−0.438,0.438]
NASA-TLX OW$_r$	38.167(18.424)	37.000(15.414)	1.167 [−4.441,6.775]	0.435	.668	0.097 [−0.343,0.535]
Time$_{total}$	4.617(1.475)	14.637(2.238)	−10.020 [−10.733,−9.307]	−29.398	<.001	−6.574 [−8.686,−4.451]
Time$_{human}$	4.617(1.475)	5.374(1.855)	−0.757 [−1.248,−0.265]	−3.224	.004	−0.721 [−1.207,−0.220]

Note. The sample size per setting is $N_p = 20$. $\Delta M = M_{\text{manual}} - M_{\text{HRC}}$. All p values are two-tailed. Cohen's d and its 95% CI are reported as provided by SPSS for paired samples, standardized using the standard deviation of the difference scores.

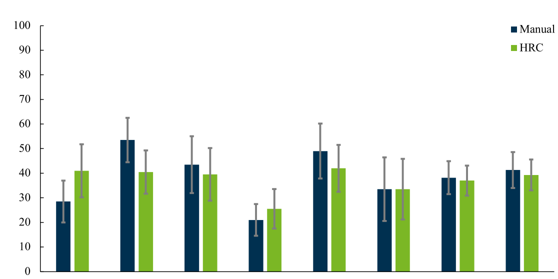
Table 6
Results of Wilcoxon signed-rank tests comparing manual and HRC conditions (**P2**).

Variable	Manual Mdn [IQR]	HRC Mdn [IQR]	Z	p	r
NASA-TLX TD	50.000 [12.500,70.000]	40.000 [20.000,50.000]	−0.528	.597	.118
NASA-TLX OP	20.000 [5.000,30.000]	30.000 [10.000,40.000]	−0.964	.335	.216
NASA-TLX OW$_w$	40.333 [20.000,61.333]	35.000 [25.833,54.167]	−0.221	.825	.049
Success rate	1.000 [1.000,1.000]	1.000 [1.000,1.000]	−1.414	.157	.316

Note. The sample size per setting is $N_p = 20$. Analyses were conducted using two-tailed Wilcoxon signed-rank tests. Mdn denotes the median; IQR represents the interquartile range (25th to 75th percentile). Effect size is calculated as $r = |Z|/\sqrt{N}$.



(a)



(b)

Fig. 8. Perceived workload assessed by NASA-TLX [69,70]. (a) Results for **P1**. (b) Results for **P2**.

the evaluation approach adopted in [74], a supplementary robustness test was conducted to quantify the valid pose estimation rate and pose-estimation accuracy of the proposed vision pipeline under different illumination conditions.

The same camera, marker detector, and hand-defined ROI procedure used in the HRC experiment were applied. The clamp-cap assembly was fixed at a known reference pose 30 cm in front of the camera. This fixed pose was used as the reference marker pose for computing

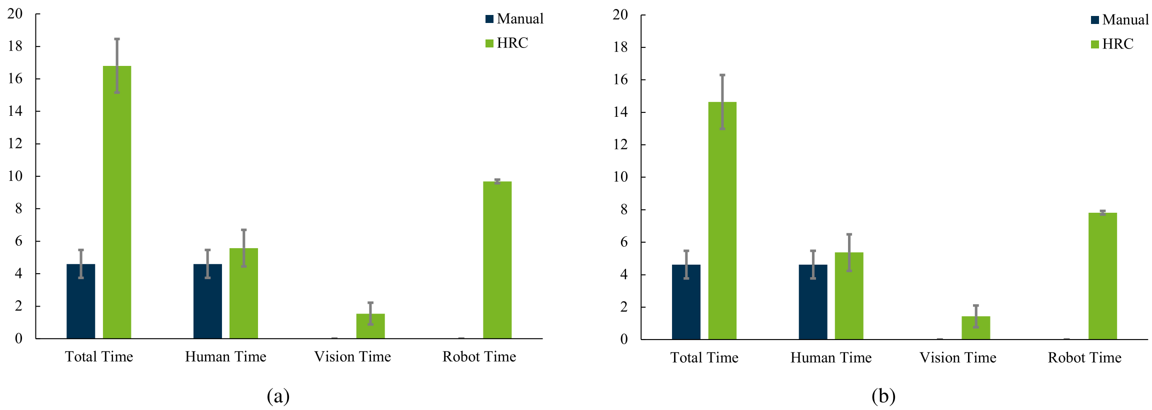


Fig. 9. Cycle time (total time) and its components (human time, vision time, and robot time) in seconds. (a) Results for **P1**. (b) Results for **P2**.

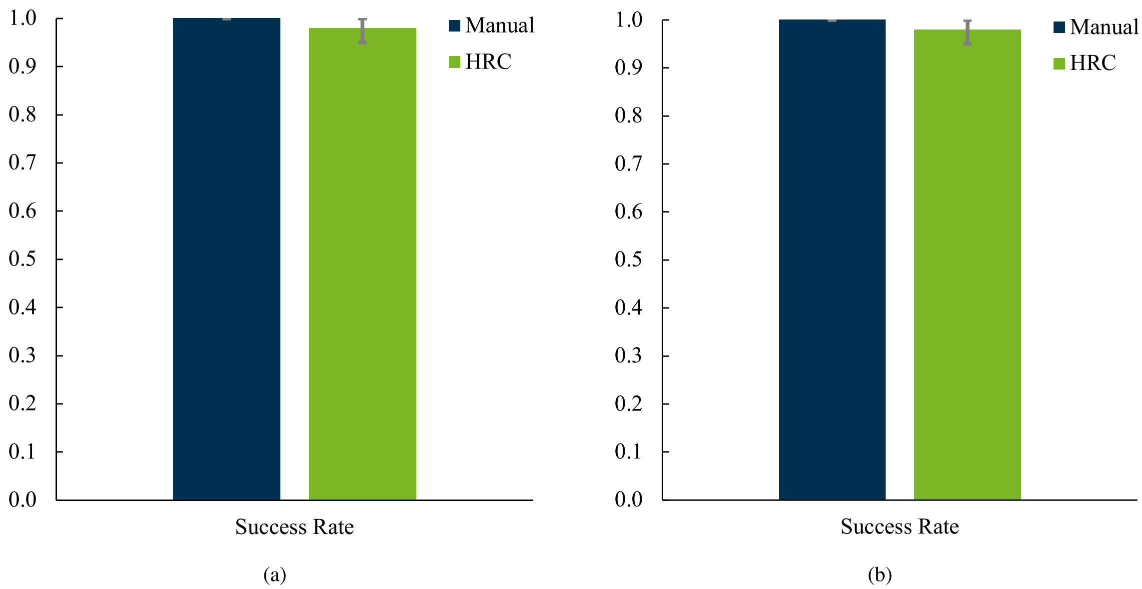


Fig. 10. Success rate of assembly. (a) Results for **P1**. (b) Results for **P2**.

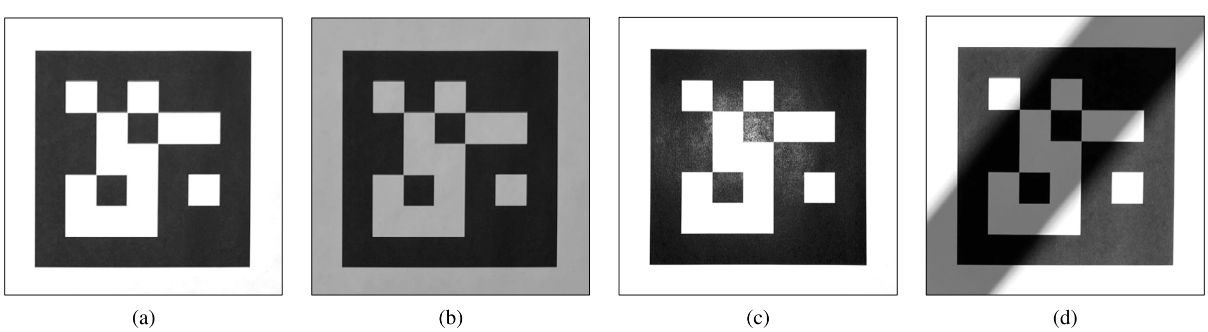


Fig. 11. The marker taken under (a) normal laboratory lighting, (b) reduced lighting, (c) perpendicular lighting, and (d) shadowed conditions.

translational and rotational deviations. Image frames were collected under four lighting conditions: normal laboratory lighting, reduced lighting, perpendicular lighting, and shadowed conditions. For each lighting condition, a 60 s video was captured at 30 fps, resulting in 1800 image frames for evaluating the vision pipeline. Fig. 11 shows representative images acquired under these lighting conditions.

For each lighting condition, the valid pose estimation rate and pose-estimation accuracy were computed. The valid pose estimation rate was defined as the proportion of frames in which the marker was detected and a 6DoF pose was estimated. Pose-estimation accuracy was evaluated using translational and rotational deviations from the fixed reference marker pose. For each valid pose estimate, the translational deviation was computed as the Euclidean distance between the estimated marker position and the reference marker position. The rotational deviation was computed as the angle of the relative rotation between the estimated marker orientation and the reference marker orientation. For each lighting condition, the mean and standard deviation (SD) of the translational and rotational deviations were then calculated across all valid pose estimates. These deviations quantify the accuracy of the estimated marker pose relative to the fixed reference pose.

Table 7
Robustness evaluation of ArUco marker detection and pose-estimation accuracy under different illumination conditions.

Lighting condition	Valid pose estimation rate	Translational deviation (mm)		Rotational deviation (°)	
		Mean	SD	Mean	SD
Normal	99.388	1.617	0.852	0.779	0.068
Reduced	99.056	1.972	0.960	0.976	0.136
Perpendicular	99.500	1.592	0.741	0.765	0.066
Shadowed	98.222	2.340	1.592	2.219	0.435

Note. Each lighting condition was evaluated using 1800 image frames captured from a 60 s video at 30 fps. The valid pose estimation rate is reported as a percentage and denotes the proportion of frames in which the marker was detected and a 6DoF pose was estimated. Translational deviation is reported in millimeters, and rotational deviation is reported in degrees. The deviation values were computed relative to the fixed reference marker pose.

As shown in Table 7, the valid pose estimation rate remained above 98% under all tested lighting conditions. The highest rate was observed under perpendicular lighting, while the lowest rate occurred under shadowed conditions. The shadowed condition also produced the largest translational and rotational deviations, indicating reduced pose-estimation accuracy when marker visibility was degraded. These results provide additional quantitative evidence that the marker-based perception pipeline performs robustly when the marker is clearly visible, while also confirming that shadowing and partial occlusion remain important limitations for practical deployment.

6. Discussion

The following interpretations differentiate between effects that are directly supported by the reported statistical results and plausible explanations derived from experimental observations. When the mechanisms underlying an effect were not directly quantified, the discussion serves to contextualize the findings rather than to establish causal relationships.

6.1. Interpretation of results with respect to hypotheses

Overall, the study provides partial support for the five hypotheses outlined in Section 4.1 and demonstrates a clear trade-off among physical ergonomics, cognitive demands, and throughput. The HRC condition consistently reduced localized hand discomfort and physical demand, maintained task reliability, but increased both mental demand and cycle time relative to manual insertion.

Localized hand discomfort was significantly lower in the HRC condition during both **P1** and **P2**, thereby supporting **H1**. This result aligns with the intended function of the proposed workflow, which transfers the repetitive, force-intensive pressing action from the operator's fingers to the pneumatic end effector, while the operator remains responsible for dexterous handling and alignment. The persistence of this effect across both environments indicates that offloading the force-intensive subtask provides a localized ergonomic benefit, even under constrained access and reduced visibility.

In **P2**, general physical discomfort did not differ significantly between conditions; therefore, **H2** was not supported. Although the mean discomfort rating was lower in the HRC condition, this difference was not statistically significant. One plausible explanation is that whole-body discomfort in the car-body environment was primarily influenced by posture- and reach-related demands, such as working in confined spaces and maintaining awkward upper-body configurations, many of which remained similar between **C1** and **C2**. These factors were not directly quantified in the present study and should be considered plausible contextual explanations rather than confirmed causal mechanisms. Thus, the intervention may have reduced localized hand strain without substantially altering the broader biomechanical demands imposed by the in-vehicle workspace.

Contrary to **H3**, overall perceived workload did not decrease in the HRC condition. The NASA-TLX [69,70] results indicate that, in both phases, physical demand was significantly lower in **C2**, whereas mental demand was significantly higher. No significant differences were observed in overall workload. These findings suggest that the ergonomic benefit of robotic assistance did not result in a reduction in overall perceived workload, but rather led to a shift in workload composition from physical effort to increased cognitive effort. Observations during the experiments suggest that the current interaction design required the operator to present the clamp-cap marker within the fingertip-defined ROI, monitor the system response, and coordinate timing and hand clearance with robot motion. These activities likely introduced additional demands for attention, anticipation, and synchronization that are not present in purely manual operation. However, these contributing factors were not directly quantified in the present study and should be interpreted as observationally informed explanations rather than confirmed causal mechanisms. Therefore, the results related to **H3** indicate that the practical effectiveness of HRC depends not only on physical task redistribution but also on the quality of collaborative interaction design.

Assembly time was longer in the HRC condition during both phases, and human time also increased in **P2**; consequently, **H4** was not supported. In **P1**, the absence of a significant difference in human time suggests that the time penalty was primarily associated with HRC system components rather than with participant actions. In both **P1** and **P2**, cycle-time decomposition showed that robot execution accounted for the largest share of the HRC cycle time, while vision processing contributed only minimally. This finding indicates that robot-side execution was the main contributor to the increase in total cycle time. However, the present timing analysis did not further decompose robot execution into sub-stages such as approach, alignment, pressing, and return to standby. Therefore, the current results do not allow quantitative identification of which robot sub-stage contributed most to the observed delay. Based on the system design, the delay may be related to reduced collaborative robot speed, additional approach and return motions, and sequential perception-motion execution. Future work should include finer robot-side timing analysis to identify the dominant sources of delay and guide cycle-time optimization. Additionally, the increase in human time in **P2** may indicate additional interaction overhead under constrained conditions. Observations during the experiments indicated repeated micro-adjustments to maintain visibility or pose estimation, as well as cautious coordination with robot movement in the restricted workspace. However, these behaviors were not directly quantified and should be interpreted with caution. Overall, the results indicate that the throughput limitation of the current HRC implementation was driven primarily by robot execution time, with additional human-side interaction overhead becoming more evident in the industrially relevant setting.

The success rate did not differ significantly between conditions in either phase, thereby supporting **H5**. In **C2**, robot execution was initiated only when the operator intentionally presented the clamp within the fingertip-defined ROI and a valid ArUco marker [19–21] pose was estimated. This interaction design appears to have facilitated robot execution stability by reducing robot activation under ambiguous perception conditions. Additionally, the fixed-size planar marker provided a consistent visual reference for 6DoF pose estimation, while ROI-constrained detection limited the marker search to the task-relevant area and reduced susceptibility to false detections from the surrounding environment. Although the collaborative mode resulted in higher mental demand and longer cycle times, it maintained overall task effectiveness.

Several failures in **C2** were associated with incorrect pose estimation, highlighting a limitation of the current perception implementation. Experimental observations indicated that partial marker occlusion and random shadowing during operator presentation were the main conditions leading to these errors. Both factors can interfere with accurate corner localization and reduce the accuracy of the estimated marker pose. In these cases, although the marker remained detectable, the estimated pose did not accurately represent the true clamp pose, which adversely affected robot alignment during insertion.

The supplementary perception robustness evaluation provides quantitative support for this interpretation. The valid pose estimation rate remained high under all tested lighting conditions, exceeding 98%, and the lowest rate was observed under shadowed conditions. The shadowed condition also produced the largest translational and rotational deviations from the fixed reference marker pose, indicating reduced pose-estimation accuracy when marker visibility was degraded. In contrast, normal, reduced, and perpendicular lighting conditions resulted in smaller deviations, suggesting more accurate pose estimation when the marker was clearly visible. These findings are consistent with the perception-related failures observed during the HRC experiment and indicate that the hand-triggered, ROI-constrained perception strategy supports reliable robot triggering and clamp insertion, but does not fully eliminate errors caused by occlusion and shadowing. Therefore, marker visibility remains a key factor for reliable robot alignment and execution in the current implementation.

In summary, the findings demonstrate that the proposed vision-based HRC framework achieved its primary ergonomic objective of reducing localized hand strain without compromising task effectiveness. However, the staged evaluation identified a clear trade-off: the physical benefits of robotic assistance were associated with increased mental demand and longer cycle times. These results underscore the necessity of jointly optimizing physical task redistribution, interaction design, perception robustness, and robot-side cycle efficiency in the development of collaborative assembly systems.

6.2. Comparison with prior work

Prior research on robotic wire harness installation has primarily focused on fully automated solutions [16,17,40,41,75]. These systems typically utilize fixed external sensing configurations and either require supplementary wire harness attachments, such as clamp caps with markers [16,40,41,75], or rely on CAD models of the wire harness bundle and its components [17]. While these methods are effective under controlled laboratory conditions, they are highly sensitive to variations in part presentation, occlusions, and workspace limitations. Furthermore, the systems described in [16,40,41,75] necessitate that all clamp caps be attached prior to executing the robotic assembly program and remain on the harness after assembly. In contrast, the approach in [17] removes the need for additional attachments but requires a fully spread wire harness bundle, a uniform background for point-cloud segmentation, and a CAD model for registration and component localization. These constraints underscore the extent to which full automation in wire harness installation depends on tightly controlled conditions that are difficult to maintain in real automotive final assembly environments.

In contrast to robot-centric automation strategies, this study introduces a collaboration-oriented framework in which the operator selects, prepares, and positions the clamp, while the robot executes the repetitive, force-intensive pressing step using vision-guided pose estimation. Rather than pursuing complete automation, the proposed human-centered method is designed for shared-workspace collaboration within the practical constraints of automotive final assembly, where variability, restricted accessibility, and operator-driven task progression are prevalent. The framework prioritizes robust task handover, predictable robotic behavior, and consistent execution of the ergonomically demanding insertion action. The objective is not to claim that collaboration is universally superior to full automation, but to demonstrate that it offers a more practical solution in scenarios where robust deployment of full automation is difficult.

A key aspect of the proposed system is the intentional integration of perception and interaction. The system employs an eye-in-hand camera, marker-based 6DoF pose estimation, and an operator-guided triggering mechanism that restricts marker detection to a fingertip-defined ROI. This approach links robot execution to a specific, task-relevant operator action, rather than relying on continuous object localization throughout the workspace. Compared to previous fully automated pipelines [16,17,40,41,75], this design reduces the need for workspace-wide component localization and enables reliable pose estimation at the moment of action, even in cluttered or spatially constrained environments. The primary contribution is not the isolated use of markers or hand-triggered activation, but their integration into a practical HRC workflow for wire harness installation, accompanied by a staged evaluation of its ergonomic and operational trade-offs under increasingly realistic conditions.

This design decision is further illustrated by the role of the clamp cap within the proposed workflow. Incorporating clamp caps introduces additional hardware and handling steps, representing a form of overhead commonly observed in marker- or fixture-assisted wire harness assembly automation [12]. Within the proposed HRC context, the clamp cap functions as a standardized planar support for fiducial marker detection and provides a consistent geometric reference for pose estimation. This enhances system observability and enables more precise robot alignment during pressing. The cap is designed for rapid attachment and repeated reuse. Unlike the pre-assembled clamp caps required in previous robot-oriented systems [16,40,41,75], the current approach eliminates the need for additional wire harness preparation before installation, and the cap can be removed immediately after insertion. Compared to the attachment-free localization method described in [17], the proposed technique does not require fully spread wire harnesses, simplified backgrounds, or CAD models. Thus, the clamp cap represents a practical compromise between operational complexity and perception reliability.

This practical distinction is further evidenced by the reported system-level outcomes. Among previous studies, only [40,75], and [16] presented experimental results for a complete clamp recognition and insertion process, and only [16] reported both average cycle time and success rate, with values of 22.500 s and 0.500, respectively. Within the proposed HRC framework, success rates of 0.980 were achieved in both laboratory and car-body environments, with average cycle times of 16.802 s and 14.637 s, respectively. Although these results are not directly comparable because of differences in system architecture, task structure, and evaluation protocols, they suggest that a collaboration-oriented approach can achieve high task reliability under less restrictive environmental assumptions. Therefore, the proposed framework supports HRC as a practically meaningful alternative for real-world wire harness installation when robustness to variability and integration into human-centered assembly workflows are prioritized.

6.3. Implications for industrial deployment

The findings provide practical guidance for the design and implementation of HRC systems in assembly tasks, especially for non-rigid object assembly within constrained environments. However, these results also identify several challenges that must be resolved to enable robust deployment of such systems in actual production settings.

6.3.1. Prioritizing force-intensive subtasks for ergonomic improvement

The reduction in localized hand discomfort observed during HRC in both experimental phases demonstrates the advantages of assigning the most force-intensive and repetitive subtasks to robotic systems. For tasks involving actions such as pressing, snapping, and pushing, targeted robotic assistance reduces localized strain while maintaining overall task effectiveness. From an industrial perspective, these results support the implementation of collaborative automation at ergonomically critical subtasks to optimize the combination of human dexterity and situational judgment with robotic repeatability and physical strength.

6.3.2. Addressing cognitive ergonomics by design

The increased cognitive demand observed in the current prototype appears to be associated with deliberate triggering, continuous monitoring of system responses, and waiting for robot execution. These contributing factors were observed during the experiments but were not directly quantified in the present study. This pattern suggests that physical assistance alone does not suffice when interactions require ongoing monitoring and coordination. Future refinements should reduce these cognitive burdens by making robot behavior more transparent and predictable. Examples include explicit ready-to-proceed states, clear motion-intent cues, concise status feedback, streamlined triggering and handover processes, shorter robot-side execution time, and pipelined task execution. In production environments, excessive monitoring and coordination requirements may decrease operator acceptance, elevate the risk of interaction errors, and constrain practical usability over extended periods. Addressing these issues is essential for both usability and deployment feasibility.

6.3.3. Enhancing throughput through robot-side optimization and workflow pipelining

The observed increase in assembly time during HRC demonstrates that throughput remains a substantial barrier to widespread implementation. Time decomposition analysis identifies robot execution as the primary contributor to the overall HRC cycle time. Consequently, productivity enhancements should focus on optimizing robot-side cycle times. Strategies include minimizing travel distances through improved workstation layouts and standby pose selection, refining approach and return trajectories, and adjusting speeds and accelerations within established safety protocols [76]. Utilizing robots with greater reach can further decrease robot operation time. The current prototype executed all subtasks sequentially. However, in production settings, workflow efficiency can be increased by pipelining tasks. For instance, overlapping perception with operator motion and coordinating preparation actions across insertions may reduce delays. In industrial applications, these improvements are critical because the collaborative process must align with takt-time requirements and broader line-balancing constraints. Implementing these strategies is expected to reduce robot-induced idle time and narrow the throughput gap between manual and collaborative assembly modes.

6.3.4. Maintaining task effectiveness and quality consistency

The consistently stable success rates observed in both manual and HRC conditions suggest that collaborative assistance can be integrated without compromising task effectiveness. This outcome is especially relevant for industrial assembly, where consistent quality is critical. The results demonstrate that workers are able to adapt to collaborative sequences and that the vision-based triggering strategy enables reliable execution, even when cognitive and temporal demands increase. Nevertheless, sustaining this level of effectiveness in actual production environments will require robust perception systems capable of handling variable lighting, occlusions, workspace limitations, and operator-specific variations in component presentation. The supplementary robustness evaluation reinforces this point by showing that marker-based perception is reliable under clear visibility conditions, but remains sensitive to shadowing and partial occlusion.

6.4. Limitations and future work

Several limitations related to the study design and the prototype system should be considered when interpreting the results and improving system performance.

6.4.1. Study design

A sample size of $N_p = 20$, the use of a convenience sampling [68], and the inclusion of participants from non-industrial settings are common in controlled HRC studies [64–67]. However, these methodological choices constrain statistical power and limit the generalizability of the findings to broader worker populations [63]. Specifically, participants were recruited from non-industrial settings rather than actual shop-floor environments. Consequently, their strategies, pacing, adaptation behaviors, and fatigue profiles may not accurately reflect those of experienced production operators in real assembly contexts.

Nevertheless, the selected participant pool was deemed appropriate for this controlled study, which primarily aimed to evaluate the ergonomic and task-level effects of the proposed HRC framework under repeatable experimental conditions. Including participants with diverse backgrounds and varying prior experience partially reflects the range of operator experience that may occur in real assembly contexts with workforce turnover. However, this approach does not eliminate the limitation concerning external validity.

Additionally, the two experimental phases were conducted in a fixed order with the same participant group to ensure safe familiarization and alignment with the TRL scale [27]. Although within-phase counterbalancing was implemented, this approach may still introduce learning or fatigue effects. Future research should employ larger and more diverse samples, include professional shop-floor workers when feasible, and consider alternative study designs, such as additional counterbalancing or separate cohorts, to better isolate learning and environmental effects and enhance the industrial relevance of the findings.

6.4.2. Task and workflow scope

The evaluated task represents a specific subtask within wire harness assembly, focusing on wire harness installation rather than the complete in-vehicle routing, connecting, and fastening process. The collaboration strategy established a clear division of labor by assigning the pressing operation to the robot and delegating all other subtasks to the human operator. Although this approach facilitates controlled comparisons, it does not capture the full spectrum of possible collaborative configurations. Future research should examine a broader range of workflow scenarios and alternative task allocations, including greater robotic involvement in positioning or routing, to assess how varying autonomy levels and responsibility distributions influence ergonomics, cognitive load, and throughput.

6.4.3. System limitations and performance trade-offs

The current prototype utilizes clamp caps with fiducial markers and a hand-landmark-defined ROI to initiate robotic actions. While this design enhances system transparency and facilitates deliberate, context-aware interaction, it also introduces additional handling steps and increases cognitive workload. The experimental evaluation was limited in scale and primarily addressed ergonomic effects and task-level performance under controlled conditions. Consequently, the study did not capture a wide range of failure cases or provide a comprehensive quantitative assessment of perception reliability across diverse operating scenarios. Several perception-related failures occurred, particularly under conditions of partial marker occlusion and shadowing. The supplementary perception robustness evaluation provides additional quantitative evidence on marker detection and pose-estimation stability under different illumination conditions. However, this evaluation was still limited to controlled tests and does not fully represent the range of lighting variation, occlusion, vibration, background clutter, and operator-presentation variability expected in production environments.

Furthermore, interpretations regarding mental demand and interaction overhead were based on experimental observations rather than direct quantitative measurement, necessitating targeted follow-up studies for validation. A more systematic analysis of perception reliability will require larger-scale experiments conducted under more varied and demanding conditions. From an industrial deployment perspective, these findings suggest that maintaining robust perception under variable visibility and presentation conditions remains a significant challenge for the current implementation.

Additionally, the HRC cycle time was primarily constrained by robot execution, with the prototype performing subtasks sequentially at reduced speed rather than in parallel. Although robot execution constituted the largest component of HRC cycle time, the present study did not quantify the relative contributions of specific robot sub-stages, such as approach, alignment, pressing, and return motion. Future work should include more detailed robot-side timing analyses to identify the dominant sources of delay and support targeted optimization of collaborative cycle time.

This limitation also presents a practical barrier to deployment in production environments, where collaborative operations must meet cycle-time and takt-time requirements. Future research should address: (i) interaction refinements to reduce monitoring and coordination overhead, such as clearer intent or state feedback and predictive cues [9]; (ii) robot-side cycle-time optimization through trajectory shortening, speed and acceleration adjustments within safety constraints, and improved end-effector actuation [53]; (iii) pipelined execution strategies that overlap perception and preparation actions to minimize idle periods [10]; and (iv) more systematic evaluation of perception reliability and failure modes under varying occlusion, lighting, and presentation conditions. In the long term, marker-less perception or the use of wire harness geometry and contextual cues may reduce reliance on auxiliary artifacts while maintaining system robustness. However, these approaches will also require additional computational resources [17,45] and dataset preparation [43,77].

6.4.4. Ecological validity and longitudinal effects

Although **P2** improved ecological validity by utilizing an industrially relevant car-body setup, the study did not fully address essential production factors, including ambient noise, multitasking, time pressure, and worker fatigue during extended shifts. These factors can substantially influence both operator behavior and system performance in actual shop-floor environments. Furthermore, the study concentrated solely on short-term outcomes. Long-term considerations, such as user adaptation, learning trajectories, trust development, acceptance, and cumulative musculoskeletal effects, were not examined. Therefore, field trials and longitudinal studies are necessary to evaluate sustained usability and productivity, to determine whether the observed increase in mental demand diminishes with user experience, and to assess the robustness of the proposed HRC framework under real production conditions.

7. Conclusion

A practical vision-based HRC framework was developed and evaluated for wire harness installation in automotive final assembly. Within this framework, the robot performs repetitive, force-intensive clamp pressing, whereas the human operator manages dexterous handling, precise positioning, and task sequencing. Marker-based pose estimation, hand-triggered robot activation, and ROI-based context awareness were integrated to facilitate deliberate and context-aware collaboration. The framework was evaluated in two phases corresponding to TRL 4–6 under both laboratory and industrially relevant conditions. The results demonstrate that the proposed HRC framework reduces localized physical discomfort and physical demand while maintaining task success rates. However, the evaluation also reveals a clear trade-off: collaborative assistance improves physical ergonomics but increases mental demand and assembly time. Further cycle-time analysis indicates that robot execution is the primary contributor to the observed increase in assembly time. Collectively, this study contributes a practical HRC framework for wire harness installation and provides an empirical characterization of the trade-off among ergonomic benefit, cognitive demand, throughput, and task effectiveness in a representative non-rigid object assembly task. The results suggest that vision-based HRC is a viable strategy in contexts where robust full automation is difficult to implement. However, broader industrial adoption will require improvements in interaction fluency, perception robustness, cognitive ergonomics, and robot-side cycle time. Future research will focus on reducing robot-side execution time, introducing pipelined execution strategies, refining interaction design, and minimizing reliance on auxiliary artifacts while maintaining robust perception and execution.

CRediT authorship contribution statement

Hao Wang: Writing – review & editing, Writing – original draft, Visualization, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Omkar Salunkhe:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation. **Annalena Hartmann:** Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis. **Sven Ekered:** Writing – review & editing, Validation, Resources. **Patrick Bründl:** Writing – review & editing, Writing – original draft, Supervision. **Jörg Franke:** Writing – review & editing, Supervision. **Johan Stahre:** Writing – review & editing, Supervision. **Björn Johansson:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Funding acquisition.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used Grammarly and OpenAI's GPT-5 in order to enhance the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the EWASS project, funded by the Swedish innovation agency, Vinnova and the strategic innovation program, Produktion2030, under grant number 2022-01279. This work was also supported by the MAXBATT project, funded by Region Västra Götaland, under grant number MRU 2024-00381. The work was conducted within Chalmers Production Area of Advance. Wiretronic AB and Volvo Car Corporation generously provided the automotive wire harnesses and the clamps used in this study. The authors gratefully acknowledge the support and contributions of all study participants.

Data availability

Data will be made available on request.

References

[1] X. Xu, T. Ji, P. Zheng, L. Wang, Human-centric manufacturing: Re-thinking, Re-justifying, and Re-envisioning, J. Manuf. Syst. 84 (2026) 259–268, http://dx.doi.org/10.1016/j.jmsy.2025.12.001.

[2] A. Giampieri, J. Ling-Chin, Z. Ma, A. Smallbone, A. Roskilly, A review of the current automotive manufacturing practice from an energy perspective, Appl. Energy 261 (2020) 114074, http://dx.doi.org/10.1016/j.apenergy.2019.114074.

[3] J. Heyn, P. Gümbel, P. Bobka, F. Dietrich, K. Dröder, Application of artificial neural networks in force-controlled automated assembly of complex shaped deformable components, Procedia CIRP 79 (2019) 131–136, http://dx.doi.org/10.1016/j.procir.2019.02.027.

[4] H. Yin, A. Varava, D. Kragic, Modeling, learning, perception, and control methods for deformable object manipulation, Sci. Robot. 6 (54) (2021) eabd8803, http://dx.doi.org/10.1126/scirobotics.abd8803.

[5] S. Makris, F. Dietrich, K. Kellens, S.J. Hu, Automated assembly of non-rigid objects, CIRP Ann 72 (2) (2023) 513–539, http://dx.doi.org/10.1016/j.cirp.2023.05.003.

[6] O. Salunkhe, W. Quadrini, H. Wang, J. Stahre, D. Romero, L. Fumagalli, D. Lämkull, Review of current status and future directions for collaborative and semi-automated automotive wire harnesses assembly, Procedia CIRP 120 (2023) 696–701, http://dx.doi.org/10.1016/j.procir.2023.09.061.

[7] W.P. Neumann, S. Winkelhaus, E.H. Grosse, C.H. Glock, Industry 4.0 and the human factor – A systems framework and analysis methodology for successful development, Int. J. Prod. Econ. 233 (2021) 107992, http://dx.doi.org/10.1016/j.ijpe.2020.107992.

[8] A. Keshvarparast, D. Battini, O. Battaia, A. Pirayesh, Collaborative robots in manufacturing and assembly systems: literature review and future research agenda, J. Intell. Manuf. 35 (5) (2024) 2065–2118, http://dx.doi.org/10.1007/s10845-023-02137-w.

[9] S. Li, P. Zheng, S. Liu, Z. Wang, X.V. Wang, L. Zheng, L. Wang, Proactive human–robot collaboration: Mutual-cognitive, predictable, and self-organising perspectives, Robot. Comput.-Integr. Manuf. 81 (2023) 102510, http://dx.doi.org/10.1016/j.rcim.2022.102510.

[10] S. Asif, T.C. Callari, F. Khan, I. Eimontaite, E.-M. Hubbard, M.S. Bahraini, P. Webb, N. Lohse, Exploring tasks and challenges in human-robot collaborative systems: A review, Robot. Comput.-Integr. Manuf. 97 (2026) 103102, http://dx.doi.org/10.1016/j.rcim.2025.103102.

[11] L. Zhou, L. Zhang, N. Konz, Computer vision techniques in manufacturing, IEEE Trans. Syst. Man Cybern.: Syst. 53 (1) (2023) 105–117, http://dx.doi.org/10.1109/TSMC.2022.3166397.

[12] H. Wang, O. Salunkhe, W. Quadrini, D. Lämkull, F. Ore, M. Despeisse, L. Fumagalli, J. Stahre, B. Johansson, A systematic literature review of computer vision applications in robotized wire harness assembly, Adv. Eng. Inform. 62 (2024) 102596, http://dx.doi.org/10.1016/j.aei.2024.102596.

[13] G.E. Navas-Reascos, D. Romero, C.A. Rodriguez, F. Guedea, J. Stahre, Wire harness assembly process supported by a collaborative robot: A case study focus on ergonomics, Robotics 11 (6) (2022) http://dx.doi.org/10.3390/robotics11060131.

[14] J. Sanchez, J.-A. Corrales, B.-C. Bouzgarrou, Y. Mezouar, Robotic manipulation and sensing of deformable objects in domestic and industrial applications: a survey, Int. J. Robot. Res. 37 (7) (2018) 688–716, http://dx.doi.org/10.1177/0278364918779698.

[15] H. Wang, O. Salunkhe, W. Quadrini, D. Lämkull, F. Ore, B. Johansson, J. Stahre, Overview of computer vision techniques in robotized wire harness assembly: Current state and future opportunities, Procedia CIRP 120 (2023) 1071–1076, http://dx.doi.org/10.1016/j.procir.2023.09.127.

[16] X. Jiang, K. mo Koo, K. Kikuchi, A. Konno, M. Uchiyama, Robotized assembly of a wire harness in a car production line, Adv. Robot. 25 (3–4) (2011) 473–489, http://dx.doi.org/10.1163/016918610X551782.

[17] M. Wnuk, M. Zürn, M. Paukner, S. Ulbrich, A. Lechler, A. Verl, Case study on localization for robotic wire harness installation, in: Advances in Automotive Production Technology – Towards Software-Defined Manufacturing and Resilient Supply Chains, 2023, pp. 333–343, http://dx.doi.org/10.1007/978-3-031-27933-1_31.

[18] M. Zürn, M. Wnuk, A. Lechler, A. Verl, Topology matching of branched deformable linear objects, in: 2023 IEEE International Conference on Robotics and Automation, ICRA, 2023, pp. 7097–7103, http://dx.doi.org/10.1109/ICRA48891.2023.10161483.

[19] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas, M. Marín-Jiménez, Automatic generation and detection of highly reliable fiducial markers under occlusion, Pattern Recognit. 47 (6) (2014) 2280–2292, http://dx.doi.org/10.1016/j.patcog.2014.01.005.

[20] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas, R. Medina-Carnicer, Generation of fiducial marker dictionaries using Mixed Integer Linear Programming, Pattern Recognit. 51 (2016) 481–491, http://dx.doi.org/10.1016/j.patcog.2015.09.023.

[21] F.J. Romero-Ramirez, R. Muñoz-Salinas, R. Medina-Carnicer, Speeded up detection of squared fiducial markers, Image Vis. Comput. 76 (2018) 38–47, http://dx.doi.org/10.1016/j.imavis.2018.05.004.

[22] S. El Zaatari, M. Marei, W. Li, Z. Usman, Cobot programming for collaborative industrial tasks: An overview, Robot. Auton. Syst. 116 (2019) 162–180, http://dx.doi.org/10.1016/j.robot.2019.03.003.

[23] Nelfiyanti, H.N.M.Z.N. Mohamed, M. Rashid, Analysis of measurement and calculation of MSD complaint of chassis assembly workers using OWAS, RULA and REBA method, Int. J. Automot. Mech. Eng. 19 (2) (2022) 9681–9692, http://dx.doi.org/10.15282/ijame.19.2.2022.05.0747.

[24] J.M. Muggleton, R. Allen, P.H. Chappell, Hand and arm injuries associated with repetitive manual work in industry: a review of disorders, risk factors and preventive measures, Ergonomics 42 (5) (1999) 714–739, http://dx.doi.org/10.1080/001401399185405.

[25] R.M. Reyes, J. de la Riva, A. Maldonado, A. Woocay, R. de la O, Association between human error and occupational accidents' contributing factors for hand injuries in the automotive manufacturing industry, Procedia Manuf. 3 (2015) 6498–6504, http://dx.doi.org/10.1016/j.promfg.2015.07.936.

[26] O. Salunkhe, J. Stahre, D. Romero, D. Li, B. Johansson, Specifying task allocation in automotive wire harness assembly stations for Human-Robot collaboration, Comput. Ind. Eng. 184 (2023) 109572, http://dx.doi.org/10.1016/j.cie.2023.109572.

[27] J.C. Mankins, Technology readiness assessments: A retrospective, Acta Astronaut. 65 (9) (2009) 1216–1223, http://dx.doi.org/10.1016/j.actaastro.2009.03.058.

[28] C. Kienle, B. Alt, F. Schneider, T. Pertlwieser, R. Jäkel, R. Rayyes, AI-based framework for robust model-based connector mating in robotic wire harness installation, in: 2025 IEEE 21st International Conference on Automation Science and Engineering, CASE, 2025, pp. 2444–2449, http://dx.doi.org/10.1109/CASE58245.2025.11164054.

[29] M. Saadat, P. Nan, Industrial applications of automatic manipulation of flexible materials, Ind. Robot.: Int. J. 29 (5) (2002) 434–442, http://dx.doi.org/10.1108/01439910210440255.

[30] N. Lv, J. Liu, Y. Jia, Dynamic modeling and control of deformable linear objects for Single-Arm and Dual-Arm robot manipulations, IEEE Trans. Robot. 38 (4) (2022) 2341–2353, http://dx.doi.org/10.1109/TRO.2021.3139838.

[31] A. Monguzzi, T. Dotti, L. Fattorelli, A.M. Zanchettin, P. Rocco, Optimal model-based path planning for the robotic manipulation of deformable linear objects, Robot. Comput.-Integr. Manuf. 92 (2025) 102891, http://dx.doi.org/10.1016/j.rcim.2024.102891.

[32] S. Javdani, S. Tandon, J. Tang, J.F. O'Brien, P. Abbeel, Modeling and perception of deformable one-dimensional objects, in: 2011 IEEE International Conference on Robotics and Automation, 2011, pp. 1607–1614, http://dx.doi.org/10.1109/ICRA.2011.5980431.

[33] A. Keipour, M. Bandari, S. Schaal, Deformable one-dimensional object detection for routing and manipulation, IEEE Robot. Autom. Lett. 7 (2) (2022) 4329–4336, http://dx.doi.org/10.1109/LRA.2022.3146920.

[34] E. Aguirre, B. Raucent, Economic comparison of wire harness assembly systems, J. Manuf. Syst. 13 (4) (1994) 276–288, http://dx.doi.org/10.1016/0278-6125(94)90035-3.

[35] Y. Cho, M. Cho, J. Park, B.-K. Han, Y.H. Lee, S.-H. Song, C. Park, D.I. Park, Strategic algorithm for cable wiring using dual arm with compliance control, Robot. Comput.-Integr. Manuf. 93 (2025) 102924, http://dx.doi.org/10.1016/j.rcim.2024.102924.

[36] M. Wnuk, C. Hinze, M. Zürn, Q. Pan, A. Lechler, A. Verl, Tracking branched deformable linear objects with structure preserved registration by branch-wise probability modification, in: 2021 27th International Conference on Mechatronics and Machine Vision in Practice (M2VIP), 2021, pp. 101–108, http://dx.doi.org/10.1109/M2VIP49856.2021.9665147.

[37] T. Toner, V. Molazadeh, M. Saez, D.M. Tilbury, K. Barton, Sequential manipulation of deformable linear object networks with endpoint pose measurements using adaptive model predictive control, in: 2024 IEEE International Conference on Robotics and Automation, ICRA, 2024, pp. 12585–12591, http://dx.doi.org/10.1109/ICRA57147.2024.10611551.

[38] A. Caporali, K. Galassi, R. Zanella, G. Palli, GNN topology representation learning for deformable multi-linear objects dual-arm robotic manipulation, IEEE Trans. Autom. Sci. Eng. 22 (2025) 14738–14751, http://dx.doi.org/10.1109/TASE.2025.3562231.

[39] H. Zhou, S. Li, Q. Lu, J. Qian, A practical solution to deformable linear object manipulation: A case study on cable harness connection, in: 2020 5th International Conference on Advanced Robotics and Mechatronics, ICARM, 2020, pp. 329–333, http://dx.doi.org/10.1109/ICARM49381.2020.9195380.

[40] K. Koo, X. Jiang, K. Kikuchi, A. Konno, M. Uchiyama, Development of a robot car wiring system, in: 2008 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 2008, pp. 862–867, http://dx.doi.org/10.1109/AIM.2008.4601774.

[41] X. Jiang, Y. Nagaoka, K. Ishii, S. Abiko, T. Tsujita, M. Uchiyama, Robotized recognition of a wire harness utilizing tracing operation, Robot. Comput.-Integr. Manuf. 34 (2015) 52–61, http://dx.doi.org/10.1016/j.rcim.2014.12.002.

[42] H.-C. Song, Y.-L. Kim, D.-H. Lee, J.-B. Song, Electric connector assembly based on vision and impedance control using cable connector-feeding system, J. Mech. Sci. Technol. 31 (12) (2017) 5997–6003, http://dx.doi.org/10.1007/s12206-017-1144-7.

[43] H. Wang, G. Urbanos Uriel, K. El-Nahass, S. Ekered, B. Johansson, Accelerating industrial vision: Systematic robot-assisted dataset preparation for object detection and pose estimation, Eng. Appl. Artif. Intell. 176 (2026) 114741, http://dx.doi.org/10.1016/j.engappai.2026.114741.

[44] T.P. Nguyen, J. Yoon, A novel vision-based method for 3D profile extraction of wire harness in robotized assembly process, J. Manuf. Syst. 61 (2021) 365–374, http://dx.doi.org/10.1016/j.jmsy.2021.10.003.

[45] H. Wang, B. Johansson, Deep learning-based connector detection for robotized assembly of automotive wire harnesses, in: 2023 IEEE 19th International Conference on Automation Science and Engineering, CASE, 2023, pp. 1–8, http://dx.doi.org/10.1109/CASE56687.2023.10260619.

[46] T. Hermansson, R. Bohlin, J.S. Carlson, R. Söderberg, Automatic assembly path planning for wiring harness installations, J. Manuf. Syst. 32 (3) (2013) 417–422, http://dx.doi.org/10.1016/j.jmsy.2013.04.006.

[47] P. Di, J. Huang, F. Chen, H. Sasaki, T. Fukuda, Hybrid vision-force guided fault tolerant robotic assembly for electric connectors, in: 2009 International Symposium on Micro-NanoMechatronics and Human Science, 2009, pp. 86–91, http://dx.doi.org/10.1109/MHS.2009.5352078.

[48] A. Hartmann, Z. Liu, S. Lamprecht, P. Bründl, J. Franke, AI-Driven multisensor quality inspection: A focus on robotic wire harness assembly, in: Advances in Production Management Systems. Cyber-Physical-Human Production Systems: Human-AI Collaboration and beyond, 2026, pp. 349–363, http://dx.doi.org/10.1007/978-3-032-03538-7_25.

[49] Y. Zhou, X. Jiang, D. Chen, Y. Guo, Y. Liu, Design of a gripper for cable assembly with integrated in-hand cable manipulation functions, in: 2021 IEEE International Conference on Robotics and Biomimetics, ROBIO, 2021, pp. 1929–1934, http://dx.doi.org/10.1109/ROBIO54168.2021.9739398.

[50] X. Zhang, Y. Domae, W. Wan, K. Harada, A closed-loop bin picking system for entangled wire harnesses using bimanual and dynamic manipulation, Robot. Comput.-Integr. Manuf. 86 (2024) 102670, http://dx.doi.org/10.1016/j.rcim.2023.102670.

[51] Y. Wang, Q. Guo, L. Zheng, B. Wang, P. Zheng, Z. Qi, LLM based autonomous agent of human-robot collaboration for aerospace wire harnessing assembly, Robot. Comput.-Integr. Manuf. 97 (2026) 103120, http://dx.doi.org/10.1016/j.rcim.2025.103120.

[52] V. Villani, F. Pini, F. Leali, C. Secchi, Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications, Mechatronics 55 (2018) 248–266, http://dx.doi.org/10.1016/j.mechatronics.2018.02.009.

[53] M. Dhanda, B.A. Rogers, S. Hall, E. Dekoninck, V. Dhokia, Reviewing human-robot collaboration in manufacturing: Opportunities and challenges in the context of industry 5.0, Robot. Comput.-Integr. Manuf. 93 (2025) 102937, http://dx.doi.org/10.1016/j.rcim.2024.102937.

[54] N. Robinson, B. Tidd, D. Campbell, D. Kulić, P. Corke, Robotic vision for human-robot interaction and collaboration: A survey and systematic review, ACM Trans. Hum.-Robot Interact. 12 (1) (2023) 1–66, http://dx.doi.org/10.1145/3570731.

[55] L. Johannsmeier, S. Haddadin, A hierarchical Human-Robot interaction-planning framework for task allocation in collaborative industrial assembly processes, IEEE Robot. Autom. Lett. 2 (1) (2017) 41–48, http://dx.doi.org/10.1109/LRA.2016.2535907.

[56] B. Sun, F. Chen, H. Sasaki, T. Fukuda, Robotic wiring harness assembly system for fault-tolerant electric connectors mating, in: 2010 International Symposium on Micro-NanoMechatronics and Human Science, 2010, pp. 202–205, http://dx.doi.org/10.1109/MHS.2010.5669533.

[57] O. Salunkhe, A. Syberfeldt, D. Romero, D. Lämkull, J. Stahre, Immersive human-robot collaboration in restricted or confined spaces, in: Advances in Production Management Systems. Production Management Systems for Volatile, Uncertain, Complex, and Ambiguous Environments, 2024, pp. 240–252, http://dx.doi.org/10.1007/978-3-031-71629-4_17.

[58] A.P. Lindvig, I. Iturrate, U. Kindler, C. Sloth, ur_rtde: An interface for controlling universal robots (UR) using the real-time data exchange (RTDE), in: 2025 IEEE/SICE International Symposium on System Integration, SII, 2025, pp. 1118–1123, http://dx.doi.org/10.1109/SII59315.2025.10871000.

[59] International Organization for Standardization, ISO 10218-1:2025 robotics — Safety requirements, 2025.

[60] G. Bradski, The OpenCV library, Dr. Dobb's J. Softw. Tools (2000).

[61] R. Tsai, R. Lenz, A new technique for fully autonomous and efficient 3D robotics hand/eye calibration, IEEE Trans. Robot. Autom. 5 (3) (1989) 345–358, http://dx.doi.org/10.1109/70.34770.

[62] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M.G. Yong, J. Lee, W.-T. Chang, W. Hua, M. Georg, M. Grundmann, MediaPipe: A framework for building perception pipelines, 2019, http://dx.doi.org/10.48550/arXiv.1906.08172.

[63] G. Hoffman, X. Zhao, A primer for conducting experiments in human–robot interaction, ACM Trans. Hum.-Robot Interact. 10 (1) (2020) 1–31, http://dx.doi.org/10.1145/3412374.

[64] B. Sadrfaridpour, Y. Wang, Collaborative assembly in hybrid manufacturing cells: An integrated framework for Human–Robot interaction, IEEE Trans. Autom. Sci. Eng. 15 (3) (2018) 1178–1192, http://dx.doi.org/10.1109/TASE.2017.2748386.

[65] C.-H. Chu, Y.-L. Liu, Augmented reality user interface design and experimental evaluation for human-robot collaborative assembly, J. Manuf. Syst. 68 (2023) 313–324, http://dx.doi.org/10.1016/j.jmsy.2023.04.007.

[66] K. Hald, M. Rehm, Usability evaluation framework for close-proximity collaboration with large industrial manipulators, in: 2024 IEEE International Conference on Robotics and Automation, ICRA, 2024, pp. 11209–11215, http://dx.doi.org/10.1109/ICRA57147.2024.10610568.

[67] M. Meregalli Falerni, V. Pomponi, H.R. Karimi, M. Lavit Nicora, L.A. Dao, M. Malosio, L. Roveda, A framework for human–robot collaboration enhanced by preference learning and ergonomics, Robot. Comput.-Integr. Manuf. 89 (2024) 102781, http://dx.doi.org/10.1016/j.rcim.2024.102781.

[68] S.J. Stratton, Population research: Convenience sampling strategies, Prehospital Disaster Med. 36 (4) (2021) 373–374, http://dx.doi.org/10.1017/S1049023X21000649.

[69] S.G. Hart, L.E. Staveland, Development of NASA-TLX (task load index): Results of empirical and theoretical research, in: Human Mental Workload, in: Advances in Psychology, Vol. 52, North-Holland, 1988, pp. 139–183, http://dx.doi.org/10.1016/S0166-4115(08)62386-9.

[70] S.G. Hart, Nasa-task load index (NASA-TLX); 20 years later, Proc. Hum. Factors Ergon. Soc. Annu. Meet. 50 (9) (2006) 904–908, http://dx.doi.org/10.1177/154193120605000909.

[71] D. Hu, D. DeTone, T. Malisiewicz, Deep ChArUco: Dark ChArUco marker pose estimation, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2019, pp. 8428–8436, http://dx.doi.org/10.1109/CVPR.2019.00863.

[72] M. Kalaitzakis, B. Cain, S. Carroll, A. Ambrosi, C. Whitehead, N. Vitzilaios, Fiducial markers for pose estimation: Overview, applications and experimental comparison of the ARTag, AprilTag, ArUco and stag markers, J. Intell. Robot. Syst. 101 (4) (2021) 71, http://dx.doi.org/10.1007/s10846-020-01307-9.

[73] J.L. Pulloquinga, D. Corrata, V. Mata, Á. Valera, M. Vallés, Experimental analysis of pose estimation based on ArUco markers, in: Innovations in Industrial Engineering III, 2024, pp. 138–149, http://dx.doi.org/10.1007/978-3-031-61582-5_12.

[74] M. Kalaitzakis, S. Carroll, A. Ambrosi, C. Whitehead, N. Vitzilaios, Experimental comparison of fiducial markers for pose estimation, in: 2020 International Conference on Unmanned Aircraft Systems, ICUAS, 2020, pp. 781–789, http://dx.doi.org/10.1109/ICUAS48674.2020.9213977.

[75] K. Koo, X. Jiang, A. Konno, M. Uchiyama, Clamp grasping and insertion task automation for automobile industry, in: 2010 IEEE/SICE International Symposium on System Integration, 2010, pp. 293–298, http://dx.doi.org/10.1109/SII.2010.5708341.

[76] M. Lagomarsino, M. Lorenzini, E. De Momi, A. Ajoudani, PRO-MIND: Proximity and reactivity optimization of robot motion to tune safety limits, human stress, and productivity in industrial settings, IEEE Trans. Robot. 41 (2025) 2067–2085, http://dx.doi.org/10.1109/TRO.2025.3547270.

[77] Y. Roh, G. Heo, S.E. Whang, A survey on data collection for machine learning: A Big Data - AI integration perspective, IEEE Trans. Knowl. Data Eng. 33 (4) (2021) 1328–1347, http://dx.doi.org/10.1109/TKDE.2019.2946162.