Integrating Computer Vision in Exosuits for Adaptive Support and Reduced Muscle Strain in Industrial Environments

Francesco Missiroli, *Graduate Student Member, IEEE*, Pietro Mazzoni, Nicola Lotti, Enrica Tricomi, *Graduate Student Member, IEEE*, Francesco Braghin, *Member, IEEE*, Loris Roveda, *Member, IEEE*, and Lorenzo Masia, *Senior Member, IEEE*

***Abstract*—Exosuits are wearable technologies that improve physical capabilities and mobility providing support during various activities. Although primarily intended for medical rehabilitation, there is growing interest in utilizing exosuits in industrial environments to prevent work-related musculoskeletal disorders (WMSDs) by ensuring continuous joints support. However, achieving synchronization between the exosuit and human motion, as well as effectively controlling interactions with the surroundings, presents ongoing challenges. The integration of computer vision techniques, particularly object recognition algorithms, can greatly assist exosuits in understanding the user's environment and adapting their behaviour accordingly. To address this issue, we have developed a control strategy for a soft exosuit that employs computer vision to collaboratively offer tailored assistance to the elbow, alleviating joint stress during interactions with objects of various natures and weights. We conducted a study to assess the effectiveness of the integrated system, which merges object recognition and gravity compensation within a built-in structure of the robotic exosuit. The findings confirmed that the suggested solution notably minimized muscle strain during dynamic activities, exhibiting a consistent correlation with the mass of the object being lifted, namely reducing by 45% and 54% respectively the Biceps activity while lifting the MW and HW compared to the 32% of the "Dynamic Arm". The intention of this contribution is to pave the way for incorporating the vision algorithm, thus enabling a more efficient interaction between the user and the exosuit itself. This includes adapting the control strategy to account for variations in environmental dynamics.**

***Index Terms*—Assistive robotics, computer vision, exosuits, embedded control, Industry 4.0.**

Manuscript received 12 July 2023; accepted 21 November 2023. Date of publication 29 November 2023; date of current version 12 December 2023. This letter was recommended for publication by Associate Editor K.-U. Kyung and Editor C. Laschi upon evaluation of the reviewers' comments. This work was supported by the Deutsche Forschungsgemeinschaft through Project HIT.Reha "Human Impedance control for Tailored Rehabilitation" under Grant N505327336. *(Corresponding author: Francesco Missiroli.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the IRB of Heidelberg University under Application No. S-311/2020, and performed in line with the Declaration of Helsinki.

Francesco Missiroli, Nicola Lotti, Enrica Tricomi, and Lorenzo Masia are with the Institut für Technische Informatik, Heidelberg University, Heidelberg 69120, Deutschland (e-mail: francesco.missiroli@ziti.uni-heidelberg.de; nicola.lotti@ziti.uni-heidelberg.de; enrica.tricomi@ziti.uni-heidelberg.de; lorenzo.masia@ziti.uni-heidelberg.de).

Pietro Mazzoni is with the Department of Electronics Information and Bioengineering, Politecnico di Milano, 20133 Milan, Italy (e-mail: pietro.mazzoni@mail.polimi.it).

Francesco Braghin is with the Department of Mechanical Engineering, Politecnico di Milano, 20156 Milan, Italy (e-mail: francesco.braghin@polimi.it).

Loris Roveda is with the Istituto Dalle Molle di studi sull'Intelligenza Artificiale, Scuola Universitaria Professionale della Svizzera Italiana, Università della Svizzera italiana, 6962 Lugano, Switzerland (e-mail: loris.roveda@idsia.ch).

This letter has supplementary downloadable material available at https://doi.org/10.1109/LRA.2023.3337693, provided by the authors.

Digital Object Identifier 10.1109/LRA.2023.3337693

I. Introduction

EXOSUITS, often referred to as soft exoskeletons, represent a category of wearable technology capable of enhancing user physical abilities, promoting mobility, and offering support during diverse activities [1]. In contrast to rigid exoskeletons that rely on stiff frames and motors for external support, exosuits harness flexible fabrics and soft actuators to replicate the body's natural motions.

In recent times, there's been an escalating interest within the industry for soft, active wearable robots that assist single or multiple upper extremity joints [2]. Even though these devices are primarily engineered for medical applications to rehabilitate motor control [3], [4], their versatility compared to their rigid counterparts means that exosuits can serve not only as a rehabilitative tool but also in industrial contexts [2], [5]. This involves providing consistent joint support to workers during their daily tasks to prevent work-related musculoskeletal disorders (WMSDs) and avoid muscle strain during the working tasks.

One of the major unsolved issues in exosuit development is the synchronization between the device and the wearer, along with controlling the interactions with the environment [6], [7]. The soft robotic suit needs to offer assistance to the user without impeding physiological motion, and adjusting the level of assistance based on the user's interactions with objects or devices around them [8]. This aspect becomes particularly important when the aim is to deliver assistance in an industrial setting where continuous interaction between the user and items or devices is the norm.

Vision plays a central role in human motion planning and interaction with the surrounding environment. For example, before and during contact and manipulation of objects, our brain utilizes visual information from our eyes to estimate the force that our hands need to exert. [9].

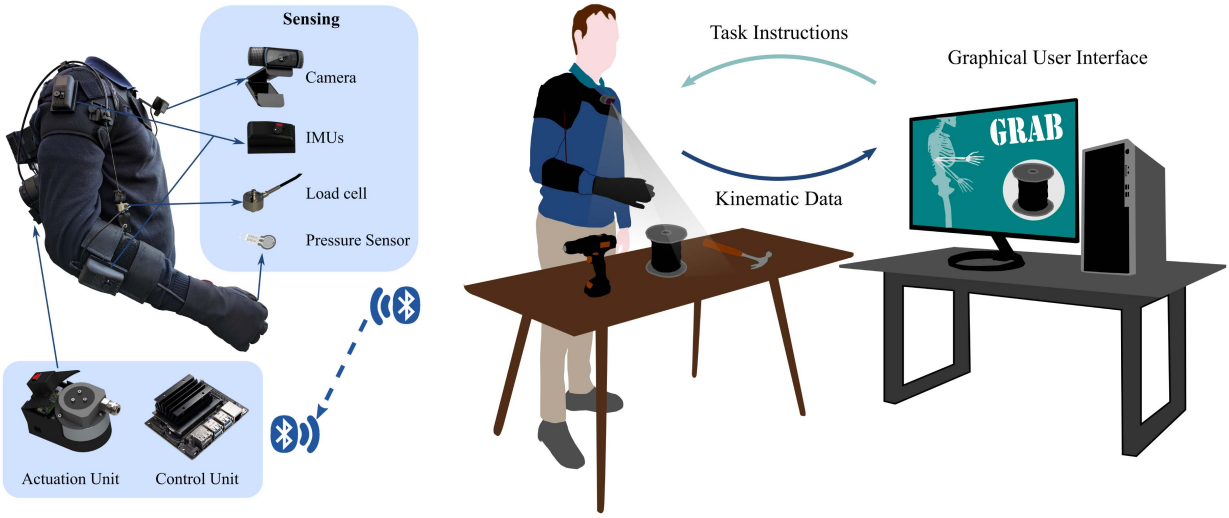


Fig. 1. Exosuit design and experimental set-up. One the left side design of the elbow exosuit: the exosuit employs fully-embedded tendon-driven mechanisms to enable elbow flexion. It consists of two key components: a lightweight orthosis and a back protector housing the actuation stages, a control unit, and a power unit. In addition, the exosuit integrates two IMUs for capturing the arm's three-dimensional orientation, a force sensor for measuring the interaction force between the user and the suit, and an RGB camera for object detection during user interactions. During the "Vision On" control mode, the assistance level adjustment is triggered by a pressure sensor located on the user's index fingertip. On the right panel, schematic representation of the experimental set-up and GUI developed for the evaluation of the control algorithm.

Bearing this principle in mind, adopting object recognition or classification algorithms for analyzing and processing visual data, a field known as computer vision, has been broadly utilized in automation [10]. This includes regulation of interactions, avoidance of unintended collisions, and more recently, fine-tuning manipulation via visual serving [11].

Despite the complexity of object recognition within the domain of computer vision, due to variations in shapes, sizes, and appearances of objects, as well as factors like lighting conditions, occlusion, and clutter [12], its potential benefits and advancements are noteworthy. In response to these challenges, a wide array of algorithmic approaches have been crafted by researchers. These span from classical computer vision methodologies [13], [14] to more recent deep learning-based techniques [15]. The evolution of these methods highlights the robust adaptability and growing capability of computer vision to overcome the challenges inherent in diverse and dynamic environments.While the application of computer vision is primarily seen in industrial robotics, there's compelling evidence to suggest that such techniques can be harnessed in assistive technology. This would provide situational awareness and discernment of environmental dynamics, subsequently allowing the assistance provided by wearable robotic devices to be tuned dynamically and adaptively [16]. By incorporating computer vision technology, exosuits can gain a better understanding of the user's surroundings and make adjustments to their behaviour accordingly [17]. For instance, the exosuit may adjust the level of assistance provided based on the objects that the user is interacting with. Currently, there have been initial attempts to incorporate computer vision into wearable robotics, either to assist upper limb joints in grasping and manipulation or lower limbs during locomotion. These efforts primarily concentrate on object recognition to facilitate the automatic grasping of desired objects by the user's hand [18], [19] or identifying stairs to adjust the assistance level provided by wearable devices based on the surrounding environment [20], [21]. Our prior investigation highlighted the continuous support abilities of exosuits designed for the elbow joint applied in the industrial environment [5]. These exosuits operate in conjunction with the user, offering the benefits of being lightweight and easily transportable. The target of our current research is to develop an occupational exosuit that actively assists the elbow during tasks that involve lifting or repetitive movements, in particular in industrial settings. Considering the wide range of movements encountered in workplaces and the various objects that users interact with, our aim is to develop a specialized algorithm that can provide customized assistance in collaboration with the user. This algorithm preserves the natural motion of the elbow by reducing strain at the joint and preventing WMSDs in situations where the arm is not loaded, as well as when interacting with objects in industrial environments.

II. Design and Control

A. Exosuit Design

The elbow exosuit (Fig. 1) comprises two main components: an embedded actuation stage using a tendon-driven principle and a customized textile harness that transfers force to the elbow joint. The device weighs 2 kg and is powered by a single battery pack (Tattu, 14.8 V, 3700 mAh, 45 C), allowing for approximately 8 hours of continuous operation. The design of the actuation prioritized lightness and portability and includes a flat brushless motor (T-Motor, AK60-6, 24 V, 6:1 planetary gear-head reduction, Cube Mars actuator, T-MOTOR, Nanchang, Jiangxi, China) that drives a pulley (35 mm diameter) around which the actuation cable is wound, as well as the control unit: this last consists in two microcontrollers running

at 100 Hz refresh frequency. The first microcontroller communicates with the sensing units via Bluetooth Low Energy (BLE, Feather nRF52 Bluefruit, Adafruit Industries, New York City, USA), while the second microcontroller (Arduino MKR 1010 WiFi, Arduino, Ivrea, Italy) is responsible for real-time control of the actuation. The sensing units, as described in Missiroli et al. [5], are composed of one microcontroller each (Feather nRF52 Bluefruit, Adafruit Industries, New York City, USA), that communicates via BLE interaction torque and the 3D kinematics of the user to the control unit. Two Inertial Measurement Units (IMUs, Bosch, BNO055, Gerlingen, Germany) detect the user's motion and reconstruct 3D arm kinematics, while a single axis load cell (ZNLBM-1, 20 kg max load, Bengbu Zhongnuo Sensor, China) connected between the actuation cable and the distal anchor point measures the interaction force between the subject and the exosuit. The transfer of force from the motor to the user is obtained by an actuation cable (Black Braided Kevlar Fiber, KT5703-06, 2.2 kN max load, Loma Linda CA, USA) that connects the user's forearm to the actuation unit via a 3D printed distal anchor point fastened on ad-hoc tailored orthosis (Sporlastic Neurolux II, Nürtingen, Germany). A second anchor point, sewn at the shoulder level, connects the actuation unit to the textile harness via a Bowden cable (Shimano SLR, 5 mm diameter, Sakai, Osaka, Japan) that absorbs resistive forces of the actuation cable (e.g., friction, backlash) and transfers them to the textile harness. To identify objects and provide an estimation of the external weight lifted, an RGB camera (Logitech C920 s PRO HD WEBCAM, Newark, CA, USA) placed on the user's chest and a dedicated embedded board (NVIDIA Jetson Nano, Santa Clara, CA, USA) running at 30 Hz was used. The decision to place the camera on the user's chest was driven by the need to closely focus on the operator's workspace, as in teleoperation and assistive robotics. This approach also provides dynamic adaptability for versatile interaction with the environment, keeping the device a standalone system.The boards communicate with each other via the Universal Serial Bus protocol. A pressure sensor (FSR 400 short, Interlink Electronics, Camarillo, California, USA) placed at the level of the index fingertip was used as a trigger to detect the interaction between the hand and the lifted object.

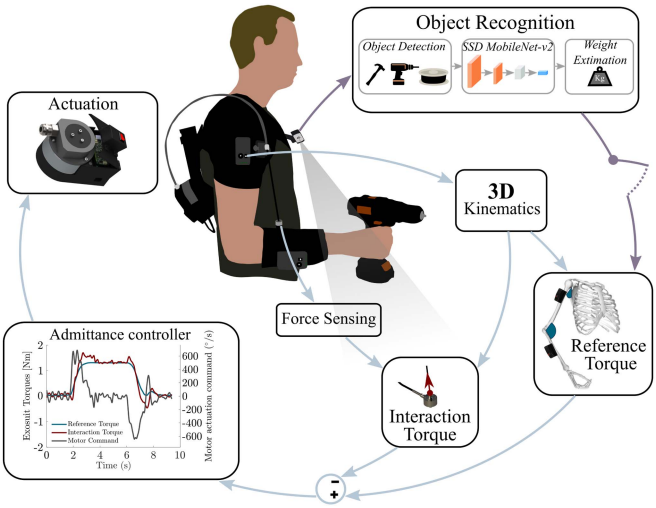


Fig. 2. Elbow exosuit real-time control framework and interaction torque. Representation of the real-time controller. The exosuit assists the wearer through the dynamic arm module by compensating for the effort required to lift the forearm against gravity and it compensates for external objects while being lifted by the user. The control framework includes a biomechanical model that is customized to the user's anthropometry with the addition of the external weight modelled as a spherical object in-built into the user's hand. The torque error computed between the reference torque and the interaction torque measured at the load cell level is then converted into a velocity input for the actuation stage via an admittance controller.

B. High-Level Control: Computer Vision Training and Parameters Optimization

For object recognition, we chose to utilize a pre-trained Convolutional Neural Network (CNN) structure known as MobileNet SSD v2. This CNN functions at a frequency of 30 Hz and runs on a dedicated board, specifically, the NVIDIA Jetson Nano.

Our training dataset comprised 1500 images collected from the authors from six distinct object classes: hammer, drill, cable spool, box, scissor, and disinfectant barrel, divided into evenly distributed subsets. These images were captured with a range of variables in mind, including object zooms, perspective angles, backgrounds, illumination, and focus. This was done to enhance the generalization of our proposed dataset and introduce variability into the network. We divided the images into three subsets using an 80/20 splitting strategy to prevent overfitting: 80% for training, 10% for validating, and 10% for testing. The pre-trained MobileNet SSD model was trained using this image dataset. We explored different combinations of batch sizes (2, 4, and 8), worker counts (1, 2, and 4), epochs (ranging from 30 to 250), and initial learning rates (0.1 or 0.01). The model with the lowest average loss, indicating the highest prediction accuracy, was selected as the best-performing model. Ultimately, we achieved optimal performance by training the model for 80 epochs, using a batch size of 4, 2 workers, a learning rate of 0.1, and an average loss of 1.085. From the output of the CNN, we selected three representative objects (hammer 0.1 Kg, drill 0.9 Kg, and cable spool 1.8 Kg, labelled as Low-Weight, Mid-Weight and High-Weight respectively along the article), chosen according to the uniform distribution in the weight change, and after linking the dataset with a lookup table we streamed the weight estimation (0.1 Kg, 0.9 Kg and 1.8 Kg respectively) and the confidence level to the Arduino board running the Dynamic Arm Control Framework.

C. Low-Level Control: Online Modulation of Robotic Assistance

The control structure (Fig. 2 left), operating on the microcontroller, carries out real-time estimation of the torque necessary during user motion and automatically tunes the assistance using estimation of objects via computer vision. The control model was built on the foundation of the "Dynamic Arm Module," extensively described by our previous work [22]: it utilizes a 3D biomechanical model of the human arm, which is customized to

the user's anthropometrics, to estimate the electromechanical assistance that the exosuit provides to the user. As earlier described, the initial controller is augmented through the incorporation of external weight estimation using computer vision techniques: when the user grasps an object, its estimated weight is used to re-estimate the dynamics of the user's arm which includes the additional weight of the grasped object. This subsequently helps to adjust the level of assistance to compensate for the added external load. Grasping is triggered by a pressure sensor located at the user's index fingertip which allows the detection of the interaction between the wearer's hand and the object, thereby activating the additional assistance in the exact moment of contact with the object. Once the reference torque (Fig. 2) is calculated using the tridimensional model of the user's arm, it is successively compared with the interaction torque measured at the end of the actuation cable via a force sensor as described in Missiroli et al. [8]: the difference between the two was fed into an admittance controller which subsequently translated the torque error into a motor velocity command and compensates for the gravity.

III. Study Protocol

In order to test the efficacy of the proposed architecture, the study involved eleven right-handed healthy participants: seven males, and four females of age $25.0[5.5]$ years (median [interquartile range(IQR)]), weight $73.0[22.0]$ Kg and height $1.83[0.13]$ m.

A. Experimental Setup

We employed a DAQ board (Quanser QPIDe, Markham, Ontario, Canada) to capture data from the control unit using a serial protocol, with a sampling frequency of 1 KHz. During the testing phase, the acquired data from the control unit provided real-time instructions to the subjects through a graphical interface (GUI, Fig. 1). For monitoring the activity of four specific muscles on the subject's right arm, we employed a wireless and multi-channel surface EMG system (Delsys Trigno, Natick MA, USA). The muscles monitored included the long head of the biceps brachii, the long head of the triceps brachii, as well as the anterior and posterior parts of the deltoid. To ensure precise electrode placement, we adhered to the SENIAM guidelines [23].

We designed the experiment in order to evaluate subjects' performance while wearing the device under three conditions to determine the physiological effects on the user in terms of change of the EMG activity and kinematics accuracy. The protocol consisted of a set of eighteen controlled movements (six repetitions for each selected object) that were performed in three different conditions, namely:

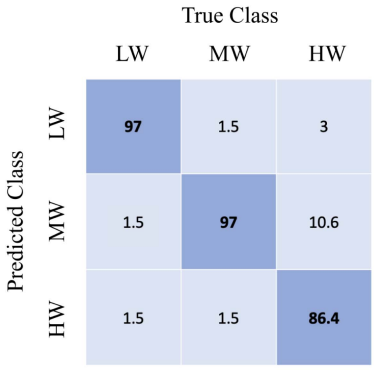
- *Exo Off:* motor off and exosuit cable slack. This was used mainly as a baseline to understand the effects of the device on the user.
- *Vision Off:* motor on, object classification and localization do not play a role in modulating the assistance. Assistance is provided only by the elbow torque computation from the *Dynamic Arm* module.
- *Vision On:* motor on, the exosuit provides adaptive assistance based on the object recognized using computer vision. Assistance is provided by the *Dynamic Arm* module which updates the motor torque with the inclusion of the extra weight placed at the level of the hand when the object is grasped.

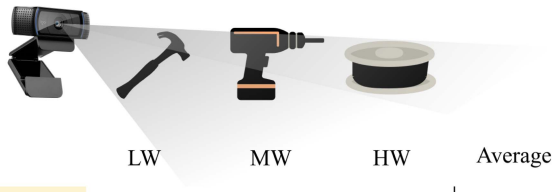
Subjects were requested to perform repetitive controlled movements, following the trajectory shown in the GUI by a phantom avatar as shown in Fig. 1. The trajectory followed by the participants comprises a series of elbow flexion-extension movements per trial with an amplitude of 90°, replicating the speed profiles of a previously published contribution [24], with an angular speed of 35°/s. A visual cue appeared on the GUI to set the start event and to tell the subject the object to lift. To allow the subject to track the avatar's movement, the GUI also showed the real-time representation of the subject's limbs overlapped with the avatar's limbs, the latter set with a transparency level of 30%.

B. Data Analysis

The effect of the computer vision algorithm on the device was quantified in terms of its effect on vision performance, muscular activity, and movement kinematics.

To evaluate the performance of the object detection model through the camera, we calculated three performance indicators: F1-Score, precision, and recall. We also reported the confusion matrix to resume multi-class classification problems, comparing the predicted and actual objects of the data set (Fig. 3). From that, we extracted the overall accuracy of the model, and assessed the impact that the vision algorithm was producing in terms of assistance delivered to the user, reporting the interaction torque computed by the model (Fig. 2). Before starting the experiments we collected maximum voluntary contraction (MVC) from the recorded muscles and used it for the EMG signal normalization during the data processing. The EMG signals were filtered offline with a fourth-order Butterworth filter (cut-off frequency 15-450 Hz), rectified and low-pass filtered at 6 Hz with a fourth-order Butterworth filter; the signals were then normalized to each participant's MVC. We used the EMG signals as a metric to quantify the effort of the users while executing the experiment. Analysis of the muscular activity was performed by comparing the traces of the four arm muscles taking into account the whole movement duration; raw EMG signals were processed offline to evaluate the Root Mean Square (RMS) as an index of activation level across tasks and conditions. The onset of the flexion and end of the elbow extension have been identified with a velocity threshold-based method [25]: first, the angle trajectories were filtered with a Savitzky-Golay filter, then the angular velocity was computed and segmented considering the 10% of velocity peak to identify the onset and offset events. Furthermore, we assessed the Co-Contraction Index (CCI) related to the elbow and shoulder joints according to the formula described in [26] to evaluate changes in the physiological motion [27]. To evaluate tracking accuracy, for each object (Low-Weight (LW), Mid-Weight (MW), High-Weight (HW)), we computed the coefficient of determination R^2 and the root mean square





	LW	MW	HW	Average
Precision	95.5%	88.9%	96.6%	93.7%
Recall	97%	97%	86.4%	93.5%
F1-score	96.2%	92.8%	91.2%	93.4%

Fig. 3. Computer vision accuracy. On the left, the confusion matrix reporting the classification accuracy per class: each class reported a mean classification accuracy above 86%. The confusion matrix shows the distribution of misclassifications across subjects. On the right, performance evaluated for the computer vision algorithm, namely precision, recall, and F1-score.

error (RMSE) by comparing the elbow trajectories measured by subjects in three different conditions (Exo Off, Vision Off, and Vision On) with the reference motion. Additionally, to determine the differences in user kinematics between the three conditions, we calculated the cross-covariance between the signals to measure the delay of movements relative to the displayed trace.

C. Statistical Analysis

We assessed the normal distribution of the measurements via a Shapiro-Wilk test with a significance level set at $\alpha = 0.05$. Repeated measures analysis of variance (rANOVA) was adopted to examine the effects of the muscle activities while performing the task. We considered the assistance type as a within-subjects factor: (*Exo Off*, *Vision Off*, and *Vision On*). Statistical significance was considered for p-values lower than 0.05; we reported the notation $F_{df1,df2}$ to indicate the degrees of freedom. Post-hoc analysis on significant main effects and interaction was performed using Bonferroni-corrected paired t-tests. Statistical analysis was conducted using Minitab (Minitab, State College, PA, USA). Reported values and measurements are presented as mean $\pm$ standard error (SE).

IV. Results

A. Vision Performance

Fig. 3 displays the evaluation of the *Computer Vision level* performance using metrics such as confusion matrix, precision, recall, and F1-score. The accuracy per class for the *low weight* (LW), *middle weight* (MW), and *high weight* (HW) is $97.0 \pm 0.8\%$, $97.0 \pm 0.7\%$, and $86.4 \pm 3.6\%$, respectively. Overall, the average precision, recall, and F1-score were found to be 93.7%, 93.5%, and 93.5%, respectively, when considering all tested objects (light, medium, and heavy). The object detection algorithm achieved an overall accuracy of 93.5%. Fig. 2-right reports the interaction torque averaged over the subjects for the "Vision Off" and "Vision On" conditions for the objects used in the experiments. We assessed that the interaction torque in the "Vision Off" condition for the three weights is similar, while there is a significant change ($F(2, 10) = 1703.77$, $p < 0.001$) with the "Vision On" condition while lifting the three different objects. In the case of the LW, the growth is $6.85 \pm 0.05\%$, for the MW it is $59.65 \pm 0.58\%$ and for the HW it is $119.84 \pm 1.25\%$.

B. Muscular Activation

In Fig. 4-left we report the EMG envelope of the two main muscles involved in the flexion/extension of the elbow (*Biceps Brachii* and *Triceps Brachii*). The Figure shows the EMG of a representative subject, averaged across trajectories and repetitions for each of the three conditions (*Exo Off*, *Vision Off*, and *Vision On*) while lifting the three performing movements with the three proposed objects (LW, MW, HW).

We observed a significant change ($p < 0.001$, $F_{4,10} = 12.18$) in the RMS of the *Biceps* activity between the "Exo Off" and "Vision Off" conditions, with values changing from $3.70 \pm 0.23\%$ to $1.88 \pm 0.14\%$ for LW (no significant difference, $p > 0.05$). For MW, the change was from $11.36 \pm 0.60\%$ to $7.70 \pm 0.61\%$ (significant difference, $p < 0.001$), and for HW, the change was from $17.20 \pm 1.00\%$ to $11.75 \pm 0.82\%$ (significant difference, $p < 0.001$). This reduction in *Biceps* activity significantly increased when introducing the computer vision algorithm in the control loop. We detected a significant change in Biceps activity between the "Exo Off" and "Vision On" conditions, with values changing from $3.70 \pm 0.23\%$ to $1.79 \pm 0.12\%$ for LW (significant difference, $p < 0.001$), from $11.36 \pm 0.60\%$ to $6.26 \pm 0.61\%$ for MW (significant difference, $p < 0.001$), and from $17.20 \pm 1.00\%$ to $7.90 \pm 0.66\%$ for HW (significant difference, $p < 0.001$). Moreover, we observed a significant change in HW Biceps activity between the "Vision Off" ($11.75 \pm 0.82\%$) and "Vision On" ($7.90 \pm 0.66\%$) conditions, with an average reduction of 33% ($p < 0.001$). However, we recorded a significant increase ($p < 0.001$, $F_{4,10} = 8.95$) in the RMS of the *Triceps* activity between the "Exo Off" and the "Vision Off" conditions for LW ($p < 0.001$). We also observed a significant increase in Triceps activity for MW and HW between the "Exo Off" and "Vision On" conditions ($p = 0.035$ for MW and $p < 0.001$ for HW). The EMG activity increased for both controllers while interacting with the LW object, going from

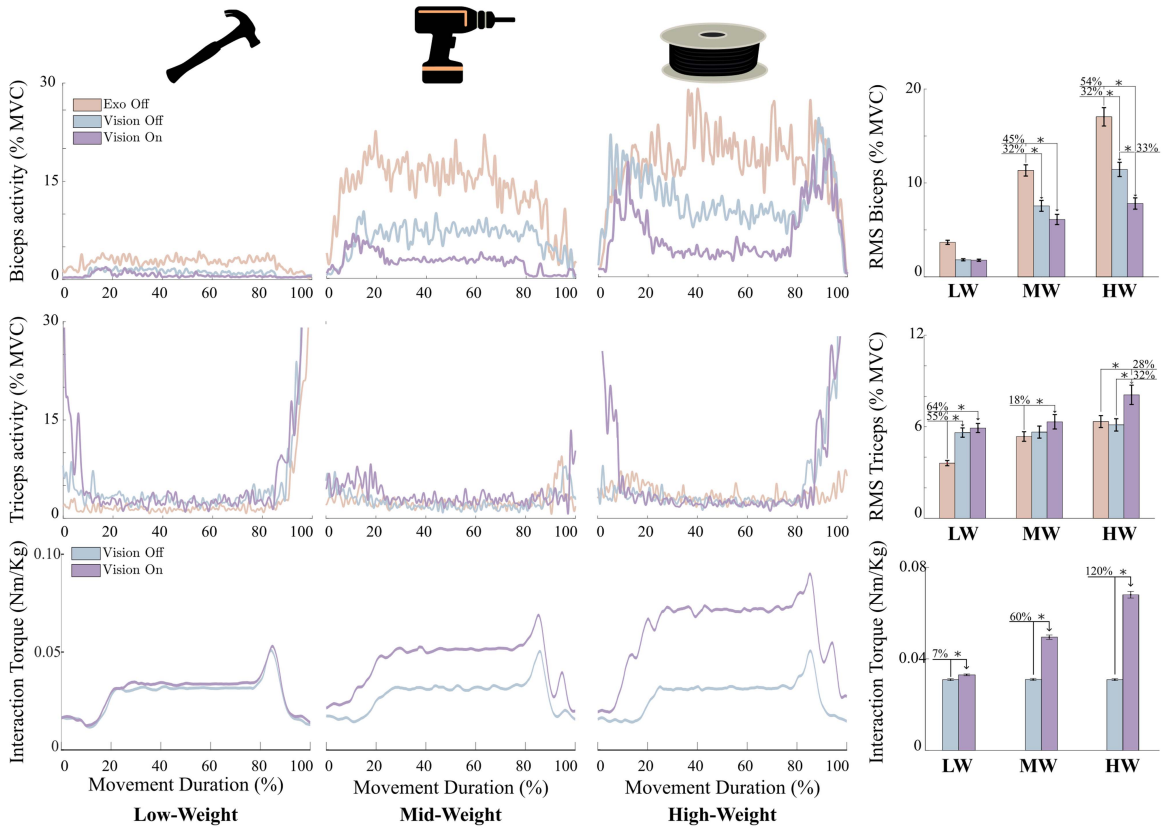


Fig. 4. EMG activity and interaction torque. On the top panel-left, time series of the *Biceps* and *Triceps* activity of a representative subject, averaged across the repetitions, performing a full movement while lifting the three proposed objects; on the top-right side RMS of the EMG activity between the conditions *Exo Off*, *Vision Off*, and *Vision On* of the main muscles involved in the elbow movements averaged across subjects; on the low panel-left, a time series of the interaction torque of a representative subject, averaged across the repetitions while lifting the three proposed objects (LW, MW, and HW) in the "Vision Off" and "Vision On" conditions; on the low panel-right, interaction torque measured by the load cell while lifting the three proposed objects (LW, MW, and HW) in the "Vision Off" and "Vision On" conditions averaged across subjects.

$3.61 \pm 0.17\%$ in the "Exo Off" condition to $5.62 \pm 0.30\%$ for "Vision Off" and $5.92 \pm 0.30\%$ for "Vision On". However, the *Triceps* activity increased only in the "Vision On" condition while interacting with the MW (from $5.36 \pm 0.32\%$ in "Exo Off" to $6.33 \pm 0.47\%$ in "Vision On") and HW (from $6.34 \pm 0.40\%$ in "Exo Off" to $8.09 \pm 0.64\%$ in "Vision On") objects. From the analysis of the *Anterior* and *Posterior Deltoids*, we did not observe any significant change in the RMS of the EMG activity. Looking at the CCI (Fig. 5), we observed a significant change ($p < 0.001$, $F_{4,10} = 21.04$) in the index for the main muscles involved in elbow motion (i.e., *Biceps* and *Triceps*). In particular, we assessed a significant reduction of the CCI between "Exo Off" and "Vision Off" ($p < 0.001$), and between "Vision Off" and "Vision On" ($p < 0.001$) with LW. The CCI changed from $65.62 \pm 2.03\%$ to $42.16 \pm 2.89\%$ and $41.56 \pm 2.96\%$ respectively, but no statistical evidence of changes with MW and HW. For *Anterior* and *Posterior Deltoids*, there was no evidence of a statistical difference between "Vision Off" ($54.81 \pm 2.46\%$) and "Vision On" ($55.01 \pm 2.21\%$) conditions compared to "Exo Off" ($54.08 \pm 1.96\%$) for LW. The same behaviour was observed for the MW ($38.00 \pm 2.20\%$, $32.25 \pm 2.52\%$, and $38.88 \pm 2.34\%$ for "Exo Off", "Vision Off" and "Vision On" respectively) and the HW ($37.50 \pm 2.26\%$,

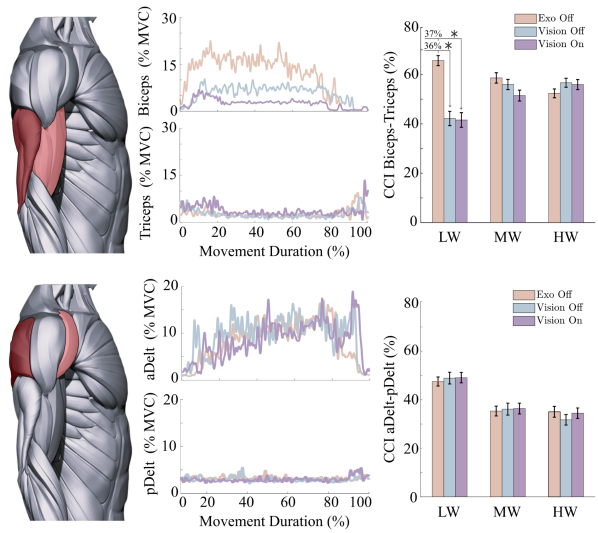


Fig. 5. Co-contraction index. On the left side the figure reports a time series of the agonist and antagonist muscles of a representative subject, averaged across the repetitions while lifting the MW; on the right side co-contraction index calculated between the two analyzed groups of muscles (*Biceps and Triceps*/*Anterior and Posterior Deltoids*) within the conditions *Exo Off*, *Vision Off*, and *Vision On* averaged across subjects computed while lifting LW, MW, and HW.

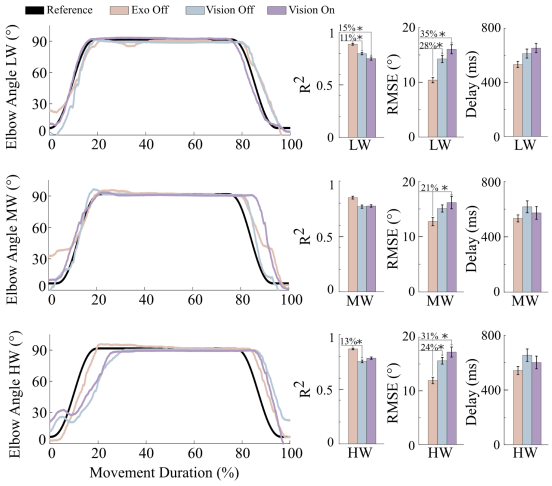


Fig. 6. Tracking accuracy. On the left side, times series of the elbow angle of a representative subject while performing the tracking task within the conditions "Exo Off," "Vision Off," and "Vision On." On the right side coefficient of determination R^2, RMSE and delay computed between the reference trajectories and the three control modalities for the three objects.

$36.80 \pm 2.38\%$, and $38.72 \pm 2.33\%$ for "Exo Off", "Vision Off" and "Vision On" respectively).

C. Movement Accuracy

Fig. 6 reports the elbow trajectories for a representative subject during the tracking task in the three analysed conditions (Exo Off, Vision Off, and Vision On) while lifting the three different loads (LW, MW and HW). We assess significant differences ($F_{4,10} = 1.32$, $p = 0.265$) in the coefficient of determination R^2 between the desired and measured trajectories, across the three conditions. In detail, we observe significant changes ($p = 0.046$ between Exo Off and Vision Off, and $p < 0.001$ between Exo Off and Vision On) in the R^2 while lifting the LW where the coefficient of determination decreased from 0.89 ± 0.01 in the "Exo Off" condition to 0.80 ± 0.01 in the "Vision Off" and 0.75 ± 0.02 in the "Vision On". This is confirmed by the RMSE ($F_{4,10} = 0.98$, $p = 0.417$), where we observe a significant change ($p < 0.001$) between the "Exo Off" (10.41 ± 0.49) and the "Vision Off" (14.28 ± 0.68), as well as ($p < 0.001$) with the "Vision On" (16.04 ± 0.83) in the LW case. This trend was repeated for the MW (12.71 ± 0.74, 15.08 ± 0.65, and 16.16 ± 1.13 for "Exo Off", "Vision Off" and "Vision On" respectively) and HW ($11.92 \pm 0.60, 15.61 \pm 0.63$, and 17.23 ± 0.94 for "Exo Off", "Vision Off" and "Vision On" respectively) objects. Looking at the delay between the reference trajectory and the subject's movement in the three conditions, we did not detect any significant change in time delay.

V. Discussion

The introduction of soft wearable devices to prevent work-related musculoskeletal disorders (WMSDs) represents a new frontier in the application of wearable technology. These devices, when integrated into fields such as logistics or assembly lines, can help alleviate stress on the user's joints. They can function either as stand-alone systems or in combination with Occupational Exoskeletons (OEs) [2], [5]. Drawing inspiration from the way visual-motor control operates in humans and its role in promoting adaptation to the surrounding environment, we improved our previously introduced control strategy to support the user's motion via an elbow exosuit [22]. The proposed strategy utilizes Computer Vision to monitor the user's surroundings and adjust the assistance provided by the exosuit accordingly. By continuously analyzing the environment, the exosuit can adapt its support to accommodate the user's needs and optimize the interaction with the surroundings. This approach leverages the power of Computer Vision to enhance the functionality and effectiveness of the elbow exosuit, promoting a more seamless and adaptive user experience.

To our knowledge, no other existing works in the literature have achieved a comparable advanced stage in this context thus far. While there are already methods to modulate the assistance of an exoskeleton or robotic suit based on muscle activity using surface electromyography (SEMG) [28], [29]. The use of SEMG has various limitations in terms of signal degradation due to user sweating, the presence of moving artefacts, and electrode detachment [30]. Consequently, especially in industrial environments, applying SEMG-based approaches is not feasible due to these reasons. To address these limitations, introducing a camera-based algorithm with computer vision in the control loop offers a promising alternative for real-time adjustment of the assistance level of a robotic suit during interaction with the surrounding environment. In this study, we present an initial example of a fully developed controller for human assistance using an elbow exosuit, accompanied by an experimental evaluation to assess its effectiveness. Our results proved that the computer vision was able to properly recognise the three chosen objects (LW, MW and HW) reporting an overall accuracy above 93.5%, with significant reduction in muscular activity once the assistance is online modulated. To elaborate, when the "Vision Off" mode is activated, the assistive torque remains constant irrespective of the weight of the object being lifted. However, when the "Vision On" mode is engaged, there is a 7% increase in assistive torque for lightweight objects, 60% for medium-weight objects, and 120% for heavy-weight objects. This variation in assistance between the "Vision On" and "Vision Off" modes has a direct impact on the Biceps muscle activity during the lifting task. Specifically, there is an average reduction of 32% in EMG activity for medium and heavy-weight objects in the "Vision Off" mode, whereas in the "Vision On" mode, muscle activity is reduced by 45% and 54% respectively compared to the "Exo Off" mode. Additionally, an increase in Triceps muscle activity was observed during the experiment. However, as highlighted in prior research by Missiroli et al. [5], this increase is likely due to the absence of assistance to the muscle during the extension phase. In upcoming research, we will concentrate on integrating a passive counteracting mechanism designed to facilitate elbow extension, thereby promoting a motion pattern that aligns more closely with a physiological one [31]. In assessing how different levels of assistance affect the user's kinematics, no significant change was observed in the delay when using the device in both "Vision Off" and "Vision On" modes compared to the "Exo

Off" mode. This indicates that the delay in the two controllers did not influence the user's reaction time during the tracking task. Nevertheless, there was a decline in the performance of the RMSE and the R^2 value when the electromechanical force of the exosuit was applied to the user. A similar adverse effect was observed in a previous study by Lotti et al. [24], where the impact was more pronounced as the speed of movement increased. This effect was limited by the actuation bandwidth and the static/Coulomb friction between the wire rope and the Bodwen sheath. It's important to note that 8 out of the 11 participants were not familiar with the device, which could have contributed to reduced accuracy in trajectory tracking. Future studies should focus on long-term observations and analyze potential learning effects related to the use of this technology. Nonetheless, there are several facets of this study that warrant further refinement. It is imperative to conduct a more comprehensive and extended study with the device to assess changes in the user's motion strategies. For example, the study should consider the users' adaptation to the mechanical assistance provided, as prolonged use of the device might yield different outcomes. We posit that movement accuracy could improve due to the wearer's ability to anticipate the exosuit's dynamics. Moreover, the experiments were carried out in a controlled environment, which fails to emulate the varied conditions of an actual industrial setting. It also does not account for the extensive array of tasks that a worker might face or the diversity of objects they might interact with. Consequently, a more extensive training dataset and a meticulous pre-calibration of the assistive technology are essential for enhancing the effectiveness of the assistance provided.

REFERENCES

[1] M. Xiloyannis et al., "Soft robotic suits: State of the art, core technologies, and open challenges," *IEEE Trans. Robot.*, vol. 38, no. 3, pp. 1343–1362, Jun. 2022.

[2] Y. M. Zhou, C. Hohimer, T. Proietti, C. T. O'Neill, and C. J. Walsh, "Kinematics-based control of an inflatable soft wearable robot for assisting the shoulder of industrial workers," *IEEE Robot. Automat. Lett.*, vol. 6, no. 2, pp. 2155–2162, Apr. 2021.

[3] T. Proietti et al., "Restoring arm function with a soft robotic wearable for individuals with amyotrophic lateral sclerosis," *Sci. Transl. Med.*, vol. 15, no. 681, 2023, Art. no. eadd1504.

[4] N. Lotti et al., "Soft robotics to enhance upper limb endurance in individuals with multiple sclerosis," *Soft Robot.*, Oct. 23, 2023, doi: 10.1089/soro.2023.0024. [Online]. Available: https://www.liebertpub.com/doi/abs/10.1089/soro.2023.0024

[5] F. Missiroli et al., "Rigid, soft, passive, and active: A hybrid occupational exoskeleton for bimanual multijoint assistance," *IEEE Robot. Automat. Lett.*, vol. 7, no. 2, pp. 2557–2564, Apr. 2022.

[6] A. Schiele, P. Letier, R. Van Der Linde, and F. Van Der Helm, "Bowden cable actuator for force-feedback exoskeletons," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2006, pp. 3599–3604.

[7] B. T. Quinlivan et al., "Assistance magnitude versus metabolic cost reductions for a tethered multiarticular soft exosuit," *Sci. Robot.*, vol. 2, no. 2, 2017, Art. no. eaah4416.

[8] F. Missiroli, N. Lotti, M. Xiloyannis, L. H. Sloot, R. Riener, and L. Masia, "Relationship between muscular activity and assistance magnitude for a myoelectric model based controlled exosuit," *Front. Robot. AI*, vol. 7, 2020, Art. no. 595844.

[9] R. Magill and D. I. Anderson, *Motor Learning and Control*. New York, USA: McGraw-Hill, 2010.

[10] H. Tian, T. Wang, Y. Liu, X. Qiao, and Y. Li, "Computer vision technology in agricultural automation—A review," *Inf. Process. Agriculture*, vol. 7, no. 1, pp. 1–19, 2020.

[11] A. Abdi, M. H. Ranjbar, and J. H. Park, "Computer vision-based path planning for robot arms in three-dimensional workspaces using Q-learning and neural networks," *Sensors*, vol. 22, no. 5, 2022, Art. no. 1697.

[12] Y. Pei, Y. Huang, Q. Zou, X. Zhang, and S. Wang, "Effects of image degradation and degradation removal to CNN-based image classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 4, pp. 1239–1253, Apr. 2021.

[13] S. Shi, Z. Wang, J. Shi, X. Wang, and H. Li, "From points to parts: 3D object detection from point cloud with part-aware and part-aggregation network," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 8, pp. 2647–2664, Aug. 2021.

[14] E. Martinez-Martin and A. P. Del Pobil, "Object detection and recognition for assistive robots: Experimentation and implementation," *IEEE Robot. Automat. Mag.*, vol. 24, no. 3, pp. 123–138, Sep. 2017.

[15] Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 11, pp. 3212–3232, Nov. 2019.

[16] P. Loncomilla, J. Ruiz-del Solar, and L. Martínez, "Object recognition using local invariant features for robotic applications: A survey," *Pattern Recognit.*, vol. 60, pp. 499–514, 2016.

[17] K. Kiguchi, Y. Kose, and Y. Hayashi, "Task-oriented perception-assist for an upper-limb powerassist exoskeleton robot," in *Proc. World Automat. Congr.*, 2010, pp. 1–6.

[18] D. Kim et al., "Eyes are faster than hands: A soft wearable robot learns user intention from the egocentric view," *Sci. Robot.*, vol. 4, no. 26, 2019, Art. no. eaav2949.

[19] J. Kuhn, J. Ringwald, M. Schappler, L. Johannsmeier, and S. Haddadin, "Towards semi-autonomous and soft-robotics enabled upper-limb exo-prosthetics: First concepts and robot-based emulation prototype," in *Proc. Int. Conf. Robot. Automat.*, 2019, pp. 9180–9186.

[20] A. G. Kurbis, B. Laschowski, and A. Mihailidis, "Stair recognition for robotic exoskeleton control using computer vision and deep learning," in *Proc. Int. Conf. Rehabil. Robot.*, 2022, pp. 1–6.

[21] E. Tricomi et al., "Environment-based assistance modulation for a hip exosuit via computer vision," *IEEE Robot. Automat. Lett.*, vol. 8, no. 5, pp. 2550–2557, May 2023.

[22] N. Lotti et al., "Myoelectric or force control? A comparative study on a soft arm exosuit," *IEEE Trans. Robot.*, vol. 38, no. 3, pp. 1363–1379, Jun. 2022.

[23] H. J. Hermens, B. Freriks, C. Disselhorst-Klug, and G. Rau, "Development of recommendations for SEMG sensors and sensor placement procedures," *J. Electromyogr. Kinesiol.*, vol. 10, no. 5, pp. 361–374, 2000.

[24] N. Lotti et al., "Intention-detection strategies for upper limb exosuits: Model-based myoelectric vs dynamic-based control," in *Proc. IEEE 8th RAS/EMBS Int. Conf. Biomed. Robot. Biomechatronics*, 2020, pp. 410–415.

[25] R. Shadmehr and F. A. Mussa-Ivaldi, "Adaptive representation of dynamics during learning of a motor task," *J. Neurosci.*, vol. 14, no. 5, pp. 3208–3224, 1994.

[26] D. A. Winter, *Biomechanics and Motor Control of Human Movement*. New York, NY, USA: Wiley, 2009.

[27] T. Bosch, J. van Eck, K. Knitel, and M. de Looze, "The effects of a passive exoskeleton on muscle activity, discomfort and endurance time in forward bending work," *Appl. Ergonom.*, vol. 54, pp. 212–217, 2016.

[28] N. Lotti et al., "Adaptive model-based myoelectric control for a soft wearable arm exosuit: A new generation of wearable robot control," *IEEE Robot. Automat. Mag.*, vol. 27, no. 1, pp. 43–53, Mar. 2020.

[29] E. Mobedi, S. Hjorth, W. Kim, E. De Momi, N. G. Tsagarakis, and A. Ajoudani, "A power-aware control strategy for an elbow effort-compensation device," *IEEE Robot. Automat. Lett.*, vol. 8, no. 7, pp. 4330–4337, Jul. 2023.

[30] R. Merletti and P. J. Parker, *Electromyography: Physiology, Engineering, and Non-Invasive Applications*, vol. 11. New York, NY, USA: Wiley, 2004.

[31] R. Chaichaowarat, J. Kinugawa, A. Seino, and K. Kosuge, "A spring-embedded planetary-geared parallel elastic actuator," in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatron.*, 2020, pp. 952–959.