Showcase 6 - Closed-loop demonstrators addressing advanced cognitive and sensorimotor functions (D3.6 – SGA3)

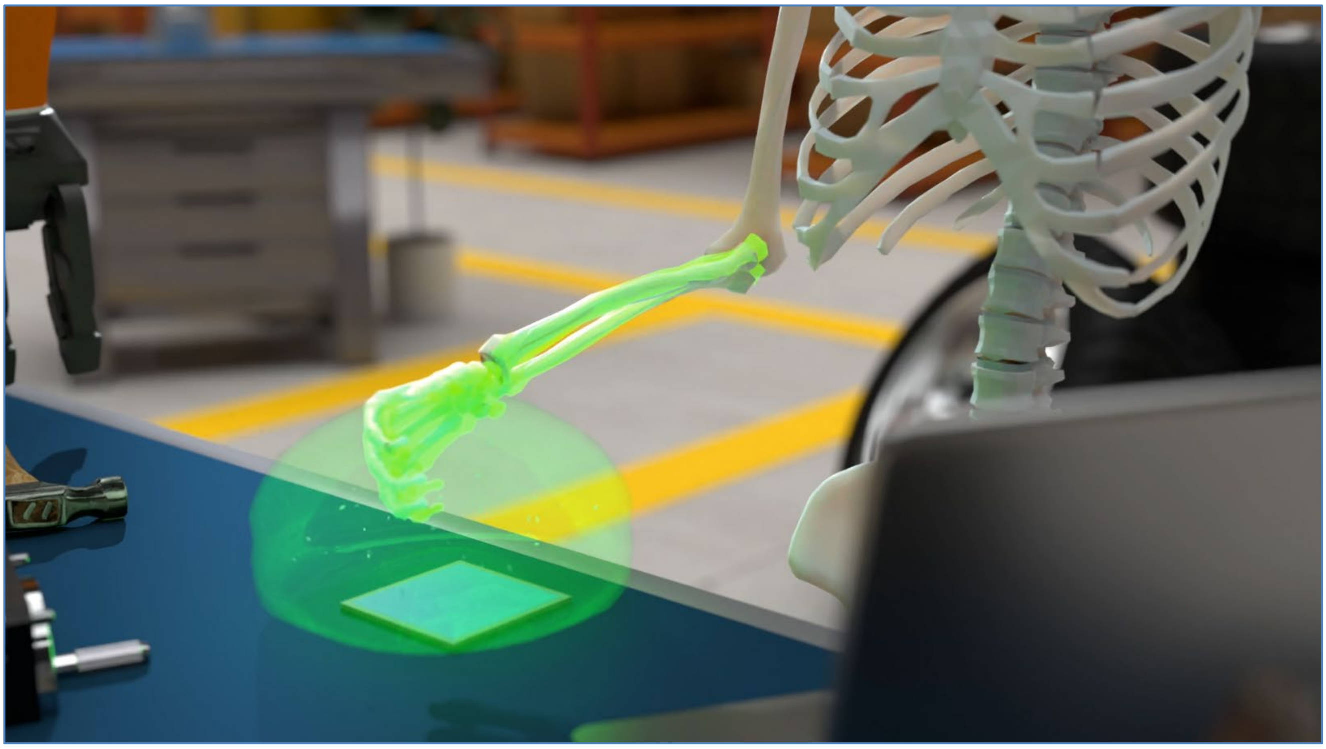


Figure 1: Digital cobotics model

Setup investigating pose estimate performance of haptic modality, a sensing surface is positioned on the workplan (blue square), its estimated detection range is materialized using a translucent volume. The estimated space occupancy of the human forearm is represented by a translucent limb, hue of which indicates estimate confidence (green in the above). When within estimation range, the estimate closely matches the actual limb, as shown here.

Project Number:	945539	Project Title:	HBP SGA3

Document Title:	Showcase 6 - Closed-loop demonstrators addressing advanced cognitive and sensori-motor functions
Document Filename:	D3.6 (D25) SGA3 M42 SUBMITTED 230929.docx
Deliverable Number:	SGA3 D3.6 (D25)
Deliverable Type:	Demonstrator
Dissemination Level:	PU = Public
Planned Delivery Date:	SGA3 M42 / 30 SEP 2023
Actual Delivery Date:	SGA3 M42 / 29 SEP 2023 (resubmitted 08 Dec 2023)
Author(s):	Yannick MOREL, UM (P117)
Compiled by:	Yannick MOREL, UM (P117)
Contributor(s):	Alban BORNET, EPFL (P134); Alice BRUEL, EPFL (P134); Ignacio ABADÍA, UGR (P66); Jörg CONRADT, KTH (P39); Jens PEDERSEN, KTH (P39); Juan Romero BERMUDEZ, KTH (P39); Egidio FALOTICO, SSSA (P49); Francesco IORI, SSSA (P49); Gojko PEROVIC, SSSA (P49); Michael ZECHMAIR, UM (P117); Raphael STOLPE, UM (P117); Tony PRESCOTT, USFD (P100)
WP QC Review:	Anita VAN OERS, UM (P117)
WP Leader / Deputy Leader Sign Off:	Rainer GOEBEL, UM (P117)
T7.4 QC Review:	N/A
Description in GA:	Release of a set of functional, closed-loop demonstrators, connecting developed cognitive architectures to relevant physical agents, building upon previously implemented sensorimotor modalities to illustrate the performance of a range of advanced cognitive tasks.
Abstract:	Deliverable D3.6 entails a number of demonstrators, composed of functional neural models, supporting a range of cognitive functions in closed loop. They are structured to address practical problems in the area of robotics and automation, specifically on the topic of human-robot interactions. The present document provides context to these developments, a summary of models implemented and results achieved by M42, and instructions on how to access the demonstrators. It further draws a number of perspectives, discussing collaborations of note that emerged from developments, applications considered, and innovation potential of specific developments.
Keywords:	Functional neural models, functional architectures, embodiment, simulation, robotics, real-world systems, real-time systems, automation and control, neuromorphic computing.
Target Users/Readers:	Robotics community, automation, control and AI community, computational neuroscientists, cognitive neuroscientists.

Table of Contents

Table of Figures

History of Changes made to this Deliverable (post Submission)

Date	Change Requested / Change Made / Other Action
29 SEP 2023	Deliverable submitted to EC
	Resubmission with specified changes requested in Review Report Main changes requested: The deliverable is acceptable in terms of scientific content, but the quality of the text is not sufficient and needs to be revised before public release. Either add details from or clearly list links to peer-reviewed publications or preprints related to the reported outcomes.
	Revised draft sent by WP to PCO. Main changes made, with indication where each change was made: Added entries in Section 2.2 to explicitly list links to peer-reviewed publications or preprints related to the reported outcomes.
	Revised version resubmitted to EC by PCO via SyGMa

1. Context

Deliverable D3.6 constitutes the final delivery of the Showcase 6 demonstrator. The Deliverable itself entails the demonstrator; that is, the demonstrator simulation model, which is made accessible (see Section 2.4 for access), and demonstration videos highlighting achievements for both the simulation model and the real-world physical demonstrator (Section 2.4). The present document accompanies the Deliverable, placing developments in the context of the Human Brain Project (HBP), outlining content of the demonstrator, relation to the state of the art, and perspectives for the technology.

The development approach for the work discussed has followed a progression from lower algorithmic abstraction level (see D3.3), towards greater abstraction, building from physical representations, to sensorimotor, to cognitive aspects (e.g. from physics simulation and rendering, to reflex behaviours, to scene understanding including pose estimation). Ambition of the work entails establishing productive collaborations across Information and Communications Technology (ICT) expertise in the project and relevant Neuroscientific expertise. The objective is twofold; identifying areas in which robotics in particular may support neuroscience (serving functions of embodiment, or physical grounding), and conversely which areas of neuroscience may help provide novel approaches to addressing established problems in ICT. Conditions of such collaboration should necessarily accommodate existing expertise within the project, and allow to address problems of genuine scientific or technological interest. This ambition had been largely fulfilled by M42, exceeding initial expectations in certain respects (with a meaningful number of identified prospective areas of fruitful collaboration not pursued due to resource limitations, see Section 3). Not all anticipated areas of collaborations however materialized in the manner initially anticipated. That is the case for instance as pertains to application of functional hierarchical planning models (investigated in WP3) to robotics planning; the work conducted on this topic in cognitive neuroscience is of excellent quality, but the fundamental tenor of functionalities investigated prevented their useful application to ICT problems. Conversely, opportunities emerged that allowed to expand productive engagement on aspects proving a natural fit between available expertise and interest from neuroscience contributors, and ability of models and neuroscientific insights provided to contribute to addressing ICT problems of relevance. This was for instance the case in the area of predictive coding for visual processing (collaboration across UM (P117) and UWE (P101) for robotic aspects, psychophysics from EPFL (P134), cognitive neuroscience contributions from UvA (P98) and UM (P117)). Similarly, collaborations with neuroscientists from UNIPIV emerged from activities related to Showcase 6, with a collaboration on embodiment of multi-area models (in NEST), with physically faithful musculoskeletal models in Bullet. Additional interdisciplinary collaborations across ICT and cognitive neuroscience emerged beyond the work in Showcase 6, including for instance the use of working memory models developed in WP3 (Kruijne, Bohte, Roelfsema, & Olivers, 2021) to support context-based behavioural adaptation for a mobile robotic system in a logistics setting (see D3.18, collaboration across cognitive neuroscientists in KNAW (P91), roboticists in UM (P117) and IIT (P135), psychologists in IIT(P135)). Below, we briefly outline the structure of the Showcase 6 demonstrator.

1.1 Final Showcase Demonstrator: Functional Architecture for Safe Human-robot Interactions

The work presented addresses a problem of practical relevance in ICT, building upon insights and expertise in neuroscience available in the HBP. The notion of productive human-robot collaboration has long since been established as a perspective of interest (Thrun, 2004). Safety concerns, however, have prevented effective deployment at scale of such technology in a real world setting (Haddadin & Croft, 2016), (Weistroffer, Keith, Bisiaux, Andriot, & Lasnier, 2022), (Rizzotti, et al., 2023). The central blocking factor is well established. Estimation of human space occupancy is typically performed using vision. However, productive interactions between human and robotic system invariably lead to situation of visual occlusion (Strazdas, Hintz and Al-Hamadi 2021); i.e. with robotic and human arm sharing an overlapping workspace, situations arise in which position of the robotic arm intercepts the line of sight from camera lens to human limb. Such situations have a substantial impact on performance of used visual systems, jeopardizing safety of the worker. Human visual processing is able to overcome partial or transient visual occlusion (Olson, 2004). The development of visual models, exploiting insights from psychophysics on the manner in which the human brain manages to mitigate issues stemming from occlusion, was conducted in the previous reporting period (as reported in D3.3). Performance of the emerging model has since been quantified, its ability to handle transient occlusion established. Additional efforts have since been invested, in a collaboration with cognitive neuroscientists in UM (P117) exploring alternate functional descriptions of predictive coding principles, investigating relative merit of different such descriptions. However, achieving safety guarantees requires additional information in situations of lasting occlusion. To address the issue, efforts were invested in implementing a haptic modality, capable of complementing the developed visual system (see Figure 1). This was conducted by UM (P117), in collaboration with USFD (Prescott, et al., 2020), (Salazar & Prescott, 2023). This haptic modality proves able to reliably estimate the pose of a human limb in situations in which it is located in close proximity of the robotic system, as discussed in the following. To exploit complementarity across modalities, we developed a multimodal visual-haptic model. This development follows a structure reminiscent of previous results from UWE (P101) on predictive coding and multimodal perception (Knowles, Stentiford, & Pearson, 2021). The collaboration with UWE (P101) in this area proved particularly helpful, substantially accelerating developments. Different multimodal implementations were ultimately investigated. Results shown in the following were obtained using a Luenberger-type nonlinear observer (Luenberger, 1971), building upon the integration of different types of neural models for either modality. In addition to work on perception, efforts were invested to expand the neural motor control model described in D3.3 (Bruel, et al., 2023), with additional musculoskeletal degrees of freedom and working trajectories learned. Similarly, the model supporting robot motion planning in the previous M21 demonstrator (Iori, et al., 2023) was extended to account for uncertain rendezvous points. Work conducted on the physical demonstrator focuses on event-based pose estimation of objects of interests, emphasizing real-time performance exploiting neuromorphic computation technology. Developments in this respects have built upon results from the previous reporting period discussed in D3.3, in particular exploiting developed software technology (Pedersen & Conradt, 2023), hardware (Bermudez, et al., 2023), and generated datasets (Turner, Pedersen, Conradt, & Nowotny, 2022) to achieve event-based object tracking (Pedersen, Singhal, & Conradt, 2023).

2. Showcase Demonstrator

In the following, we outline developments contributing to the M42 Showcase 6 demonstrator, discuss their relation to the state of the art in their respective areas, and relevance to the EBRAINS RI.

2.1 Specifications of the M42 Demonstration

Functional models included within the demonstrator are hereafter presented, distinguishing models supporting motor, planning, and perception functions.

Motor control: Activities in the area have investigated the complementarity between cerebellar motor learning and spinal cord circuitry for direct and fast muscle control. The cerebellar model is constituted of a Spiking Neural Network (SNN) equipped with synaptic plasticity at the connection between granule and Purkinje cells (work by UGR (P66)). The spinal model is equipped with stretch reflex and reciprocal inhibition (EPFL (P134)). Both cerebellar and spinal models are biologically plausible. Structure of this neural cerebello-spinal model is illustrated by Figure 2 (right), it was applied to control movements of a simplified two Degree of Freedom (DoF) model of human upper limb in the M21 demonstrator, with simplified one DoF elbow and shoulder joints (resulting limb movements in the vertical plane). The model was extended to control movements of a third DoF included in the shoulder, allowing to expand the workspace of the human limb in the demonstrator. The extension entails expansion of the neural models to account for the larger workspace. The resulting neuro-musculoskeletal model was trained to support additional trajectories for human-robot interactions, as discussed in the demonstration paragraph. In addition, this same neuro-musculoskeletal model is being adapted by UGR (P66) to develop novel active-compliance adaptive motor controller for rigid robotic systems. In complement to this neural approach to motor control, a system theoretical approach was pursued (work performed by UM (P117)). This work allowed to derive control laws defining, for a given observed external motor behaviour, corresponding muscle fibre activation profiles (Stolpe & Morel, 2023), and was extended to account for spinal dynamics. The result is of interest to neuroscientists in that it provides formal (sufficient) conditions on descending signals to support externally observed motor behaviour, helpful in further constraining functional models of motor loops (beyond physical grounding, towards *dynamical grounding*). In the context of the Showcase demonstrator, the approach allows to produce descending signals allowing to track a broad range of desired joint angle trajectories. The approach does not share the biological plausibility of the above neural controllers, but allows to straightforwardly control a broader range of human limb movements.

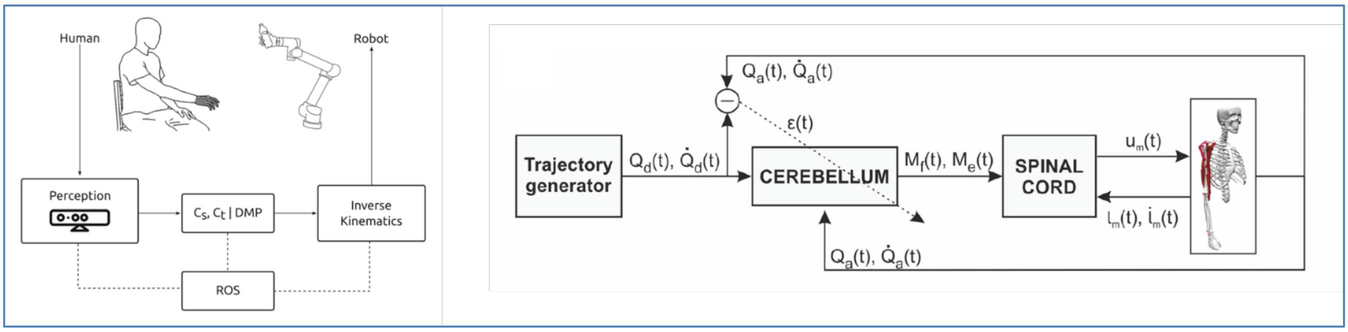


Figure 2: Block diagrams

Describing the functional structure of the planning (left) (Iori, et al., 2023) and human limb motor control (right) (Bruel, et al., 2023) algorithms.

Planning: The planning result (Iori, et al., 2023) included in the M21 demonstrator (D3.3) relied on the assumption of knowledge of an estimate of the handover location. The method used to generate online trajectories relied on Dynamic Movement Primitives (DMP). Exploiting this framework, two parametrized coupling terms were used to modulate speed of the robot's trajectory depending on the distance of the human hand from the target location and its derivative. Results of experimental tests of the developed planning technology showed a statistically significant preference from the human actor in terms of perceived satisfaction, comfort, and safety, in a comparison with alternative solutions. A simplified block diagram provides a streamlined overview of the planner's structure in Figure 2 (left). In the result in (Iori, et al., 2023), tuning of the algorithm's parameters was performed by the designer. However, preferences in terms of reactiveness of the robotic partner is expected to vary across human participants and use-case scenarios. To account for such variability, we extended the model to include a framework, based on Bayesian Optimization, for Preference Learning (PL) in adaptive handovers. Following this inclusion, a study was conducted with sixteen participants across different handover scenarios to validate accuracy of the PL algorithm, and qualitatively assess human preferences in human-robot interactions. The corresponding manuscript is currently under review. In addition, the planner in (Iori, et al., 2023) relied on knowledge of the location for the handover. During the reporting period, efforts were invested to extend the framework and relax such knowledge assumptions, allowing to address completely unstructured interactions; i.e. the robot has no a priori knowledge of the manner in which the handover will occur. To preserve online adaptation capabilities, a novel scheme was implemented to remove any dependence on a goal location. This choice allowed us to decouple the problem of *where to go* (position estimation), from that of *whether to go* (intent detection). The approach employed to pursue detection of intent builds upon notions of anticipatory control. Two networks are trained to predict both the short-term future trajectory of the human hand, as well as the direction in which the human hand should move for a handover. Monitoring inconsistencies across predictions allows the robot to infer to what extent the human is intending to go ahead with the handover. The approach was applied to allow the planned trajectory to converge directly to the predicted human hand position, rather than to a pre-defined target location, as was the case in (Iori, et al., 2023).

Perception: In the previous reporting period, EPFL (P134) developed in collaboration with UM (P117), a visual segmentation model expanding upon PredNet (Lotter, Kreiman, & Cox, 2016). Modifications of note entailed the inclusion of axonal delays, promoting emergence of temporal integration of information as in (Hogendoorn & Burkitt, 2019), the replacement of the Long Short-Term Memory units in (Lotter, Kreiman, & Cox, 2016) with horizontally Gated Recurrent Units (hGRU) to promote spatial integration, as in (Linsley, Kim, Veerabadran, Windolf, & Serre, 2018), and the addition of decoding in an approach adapted from (Long, Shelhamer, & Darrell, 2015), to reconstruct segmentation masks from representational information in the predictive coding model. Structure of the model is shown in Figure 3, an example of segmentation result, for the human skeleton model used in the simulation demonstrator, is provided in Figure 4. This model is integrated within the demonstrator's functional architecture. Its performance compares favourably to that of alternate visual segmentation models found in the literature in situations of occlusion (see Section 2.2). However, as previously mentioned, in situations of lasting occlusion, a visual modality alone is unable to provide reliable estimate of space occupancy of the human in the scene, motivating the inclusion of additional means of perception.

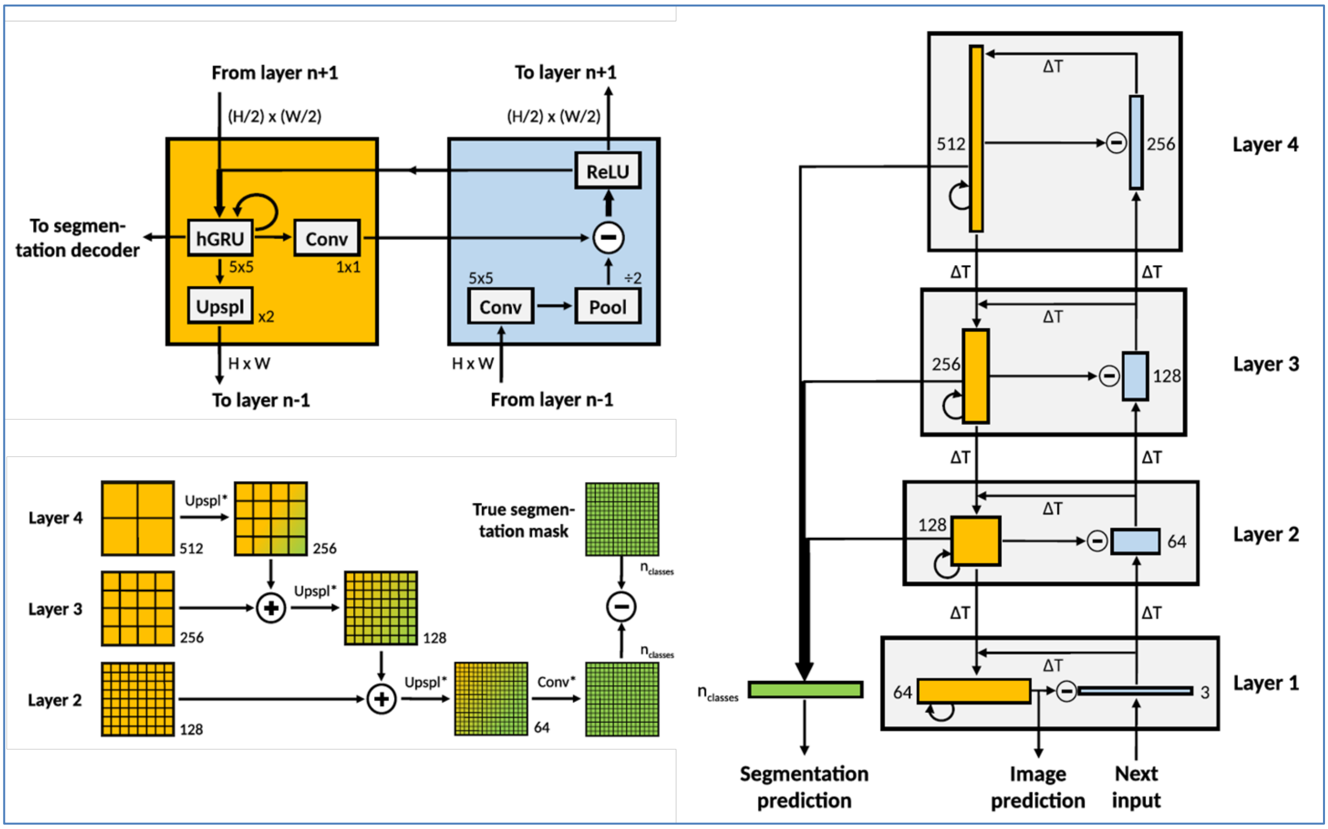


Figure 3: Structural overview of ProcNet

A robust-to-occlusion visual segmentation model based on PredNet (Lotter, Kreiman, & Cox, 2016); overall structure (right) is composed of four processing stages at which representations, constructed from latent information, are compared to afferent processed frame information. Detail of representation and error layers at each processing stage (top left), and representation information decoding into segmentation masks (bottom left).

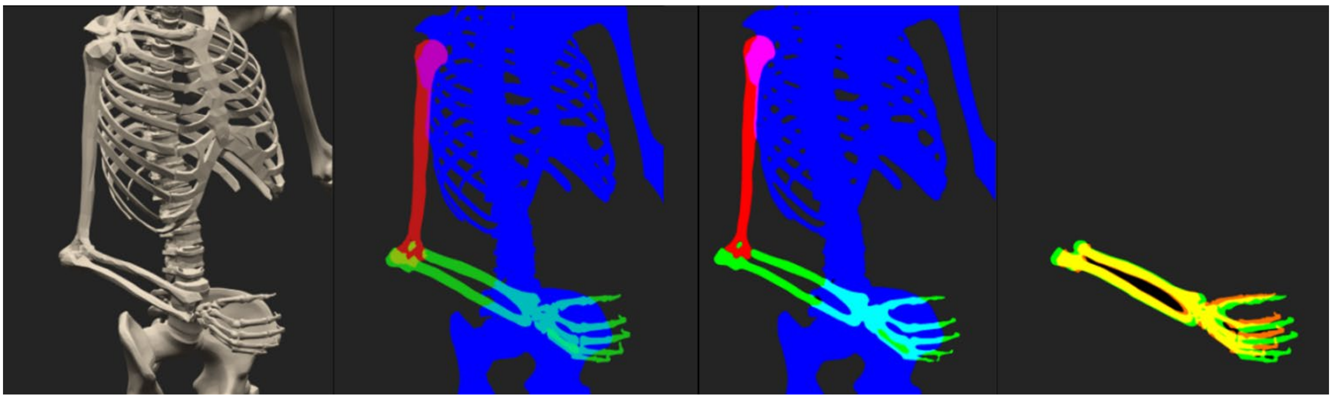


Figure 4: Visual segmentation and pose estimation of forearm using ProcNet

Model providing camera frame input (far left), ground truth segmentation masks (middle left), estimated segmentation masks produced by ProcNet (middle right), comparison of the segmentation mask of skeletal forearm for ground truth and pose estimate (far right).

We implemented a haptic modality, the technology of which was freely adapted from that described in (Schlegl, Kröger, Gaschler, Khatib, & Zangl, 2013). Sensing surfaces produce a signal, the magnitude of which reflects proximity and capacitance of objects in their immediate surroundings (decimetric detection range). A single such sensor allows to detect presence and to provide a rough estimate of the degree of proximity of a capacitive object (e.g. a human limb). Several non-collocated measures allow to infer directionality of the object, information that may be used to implement following or avoidance schemes. A greater number of measures provides yet richer information about the system's surroundings.

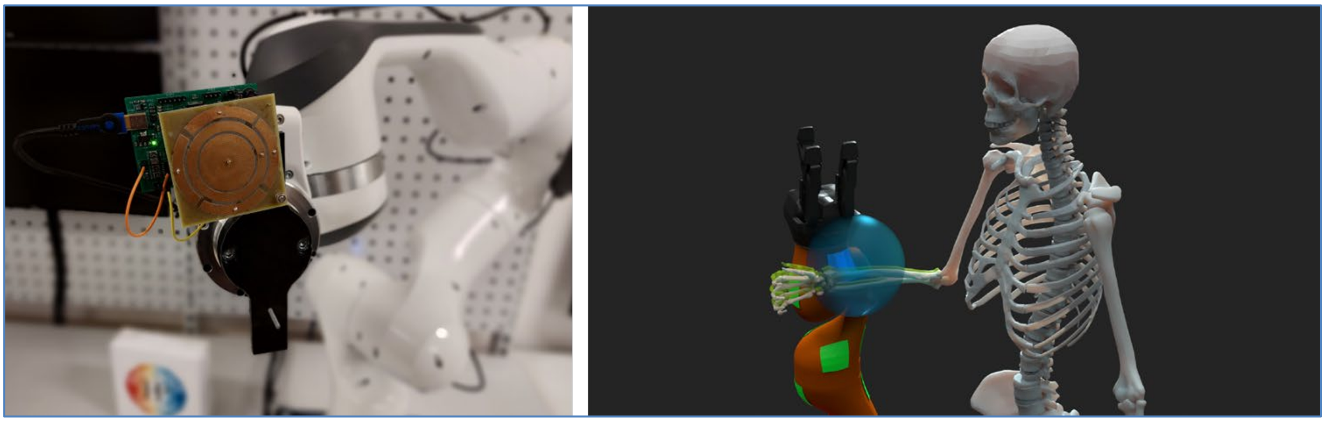


Figure 5: Capacitive sensing surfaces

Attached to the end-effector of a Franka Emika robotic arm (left), sensing patches attached to the surface of a Kuka iiwa in a simulation model, with detection volume of the top-most sensor shown in blue (right).

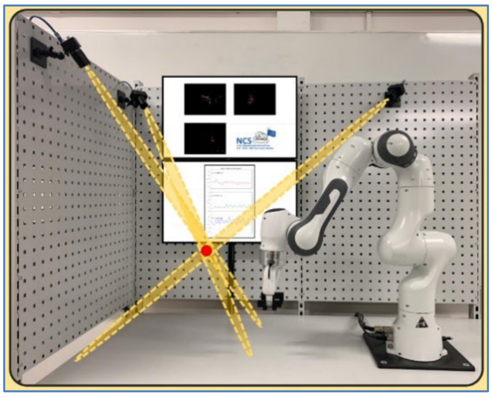
Within the previous reporting period, simulations using simplified digital models of such sensors allowed to establish utility of such a haptic modality, in particular demonstrating its complementarity with the previously developed visual perception scheme. Specifically, in situations of sustained occlusion, the (occluded) volume of space of interest is located in close proximity to the surface of the robotic arm. The availability of information descriptive of space occupancy in that volume of space may be exploited to complement information produced by the vision system, and help mitigate impact of visual occlusion. This perspective motivated the investment of efforts to implement a physical prototype of the aforementioned haptic sensor. The corresponding work was performed by UM (P117), with support from KTH (P39), which hosted developments and lent the Showcase 6 physical setup to support proceedings (see Figure 5, left). In particular, the implemented prototype allowed collection of an extensive data set, used to identify a physically faithful digital model of the modality. Development of such a model in the absence of experimental data (as was done in the previous reporting period) constitutes a problematic proposition. In particular, if the physics underlying collected measures are well described, their resolution for arbitrary electrode shapes (see sensing surfaces in Figure 5, left), metal, thickness, and surface condition, as well as for the considered spatial distribution of volumetric capacitance density of the detected object, cannot be straightforwardly resolved in closed form. It instead is typically approached using finite elements, a solution with entails meaningful approximations. Such an approximate model is sufficient to give rough indications of performance (as investigated in the previous period). It however typically proves insufficient for physically faithful simulation. The identified, digital sensor model instead proved capable of faithfully restituting results observed experimentally, and is included in the Showcase 6 simulation demonstrator. Development of models exploiting sensor measures was conducted by UM (P117) in collaboration with USFD, benefiting from their expertise in the area (Prescott, et al., 2020), (Salazar & Prescott, 2023). We were able, relying on information collected using a five-electrode setup as that shown in Figure 5 (left), to close the robotic arm's control loop and implement following and avoidance behaviours (Zechmair & Morel, Active Electric Perception-based Haptic Modality with Applications to Robotics, 2023), with consistent performance across simulation and experiments. In the presence of a sufficient number of measures, a neural model is able to learn to estimate the pose of an object present in proximity of the sensor. Such a pose estimate is shown in Figure 1, where the estimated pose is represented using a green translucent limb, and is obtained using a sensing patch of ten-by-ten electrodes. Performance in terms of accuracy of pose estimates are further discussed in Section 2.2. Detection range is of the order of 15cm; more specifically, the processing of pose estimate validation sets shows an ability for the ten-by-ten sensor in Figure 1 to estimate the relative position of a human limb with centimetric accuracy at a range of about 15cm. The general perception strategy followed consists in exploiting the aforementioned visual segmentation model to gain an understanding of the overall scene configuration, in particular of the situation of the human within this scene and of the corresponding expected space occupancy in proximity of the robotic system. Then, capacitive sensing surfaces, affixed onto carefully selected locations on the robotic system (emphasizing the last segment and end-effector, expected to come in closest proximity to the collaborating human worker), support detection of human presence in close proximity. Measures obtained are exploited using a multimodal perception approach.

Figure 6: Multimodal pose estimation of a skeletal limb

The model is placed in a purpose-specific setup designed to benchmark combined robustness to visual occlusion (right), it is also integrated within the robotic scene used in Showcase 6 (left).

This multimodal perception approach (see Figure 6 for an illustration of results) shares remarkable similarities with that considered in (Knowles, Stentiford, & Pearson, 2021), and active support from collaborators in UWE (P101) greatly accelerated developments. Use of different multimodal approaches were investigated, including that of MultiPredNet. The technique implemented within Showcase 6 exploits a nonlinear observer, which weighs contributions from both modalities using specific associated confidence levels (i.e. the capacitive sensor's signal-over-noise-ratio in the estimated relative pose of the detected object, and coherence of the segmentation masks for the pose estimated and that produced by ProcNet).

Complementary to the above-described simulated interactive demonstrator, KTH (P39) and collaborators from T5.8 (UMAN (P63)) and T5.10 (UoS (P106)) implemented a real-time, physical human-robot co-working setup, shown in Figure 7 (left). Our system consists of (1) multiple event cameras; (2) online-trained spiking neural networks for object identification, spatial localization, sensor fusion, and motor control; and (3) an actuated 7-DOF robotic arm to interact with tools from a human co-worker. The use of event-based vision sensors with continuous spiking output allows reacting to object motion within milliseconds. Our multi-layer spiking neuron network model (SNN, green in Figure 7, right) implements convolutional receptive fields to track object positions with similar performance as video cameras and non-spiking artificial neural networks achieve. Tracking information from three independent observations (multiple event cameras, each reporting in 2D) is fused into a combined estimate of the object's 3D position (Sensor Fusion, yellow in Figure 7, right). Finally, the robot arm is controlled to follow a detected tool in a safe distance. All spiking neuronal computation is executed in real-time, partially on GPUs and partially on neuromorphic hardware (SpiNNaker).



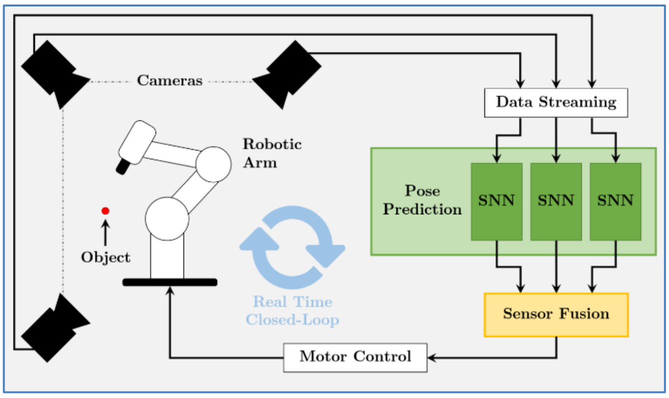


Figure 7: Human-Robot Co-working scenario

Physical demonstrator (left); control structure (right).

2.2 Scientific and Technological Problems Addressed

Hereafter, we outline the relationship to the state of the art of developments contributing to Showcase 6, distinguishing contributions to the ICT and neuroscientific literatures.

Motor control: The main question addressed by the developed cerebello-spinal model concerns the functional complementarity between described neural circuits. Motivation for approaching the modelling problem from a functional perspective stems from the notion that requiring the model to support its expected function (oftentimes in a manner that remains approximative) brings to bear additional informative constraints onto the model, which may prove helpful in ascertaining merit of assumptions underlying model construction. There exists a rich literature on functional cerebellar models, which has demonstrated merit of the approach by advancing our understanding of the role of cerebellar circuits and modules in motor functions, and in other areas the cerebellum contributes to. However, the broad majority of such functional models bypasses spinal circuitry entirely before acting on muscle models (if not on more abstract physical objects). That the spinal cord plays a substantial functional role in motor control is well established. In practice, omitting the corresponding circuitry meaningfully alters the nature of the control problem that the cerebellum model is made to address, to an extent that the particular merit of insights reached may come under question. Inclusion of a functional spinal model in the loop then certainly supports a stronger, more functionally faithful to biology, type of functional constraint. Benefits withdrawn from this inclusion were readily apparent from results achieved here; specifically, cerebellar motor learning was measurably facilitated. This manifested by more rapid cerebellar motor learning (i.e. kinematic stability reached faster), and simplified cerebellar synaptic adaptation (plastic cerebellar layer develops simpler synaptic distributions). The spinal model, in addition, played a key role in mediating cocontraction and mitigating impact of external perturbations (reducing functional burden on the cerebellar model). Detail of the findings is reported in (Bruel, et al., 2023), which is currently being reviewed for publication in PLOS Computational Biology. In complement to this contribution to the neuroscientific literature, the same functional model is being adapted to help develop a novel generation of active compliance controller for robotic systems, implementing a set of virtual muscle actuators for robotic systems. Though comparable approaches can be found in the literature, the method developed here provides an alternative allowing to exploit benefits afforded by a biologically-grounded neural approach (in particular in terms of disturbance rejection, plastic adaptation, energy minimization). Finally, the integration of the neuro-musculoskeletal model within the demonstrator also provides a setup allowing to investigate physical interactions (i.e. exchange of efforts) between robotic and human arm, accounting for kinematics (skeletal system), dynamics (muscle-tendon complexes), but also a range of human reflexes (spinal cord) and motor behaviours, which extends well beyond comparable models found in the literature on cobotics (typically avoiding consideration of effort exchanges, see the discussion in (Cherubini, Passama, Crosnier, Lasnier, & Fraisse, 2016)). At a practical level, maintaining a model of the human worker's relevant muscle-tendon complexes allows to monitor muscle and tendon strain applied by the robotic system, and inform robotic movements accordingly, supporting the development of safe-by-design robotic controllers using control barrier functions (Ames, et al., 2019). The work conducted on system-theoretical motor control (Stolpe & Morel, 2023) has direct contributions to the control theory literature, in particular addressing the muscle recruitment problem in an explicit manner, and expending on the state of the art in the area by accounting for spinal dynamics (see the discussion in (Stolpe & Morel, 2023)). The work has direct relevance to spinal stimulation technology (Lorach, et al., 2023), allowing to exploit patient-specific knowledge to inform stimulation. It is being exploited to help investigate, in closed-loop, merit of different functional models of spinal pathways. The derived conditions on descending signals can be employed to further constrain functional motor loop models, supporting a degree of *dynamical grounding*. The latter aspect is in particular investigated in active collaboration with USFD (from Task 3.1) in the context of their current investigations in the area (Prescott & Wilson, 2023).

<u>Links to peer-reviewed publications or preprints related to the reported outcomes</u>: The work on functional complementarity is under review for PLOS Computational Biology, a preprint is available on bioRxiv (doi.org/10.1101/2023.03.08.531839),

[1] Bruel, Alice, Ignacio Abadìa, Thibault Collin, Icare Sakr, Henri Lorach, Niceto Luque, Eduardo Ros, and Auke Ijspeert. "The spinal cord facilitates cerebellar upper limb motor learning and control; inputs from neuromusculoskeletal simulation." bioRxiv (2023): 2023-03.

The work on the system theoretical perspective on motor control was presented at the 2023 IEEE American Control Conference, its extension is under review for inclusion in next year's proceedings (a preprint is available from Zenodo, doi.org/10.5281/zenodo.10064098),

[2] Stolpe, Raphael, and Yannick Morel. "Model-Based Nonlinear Control of a Class of Musculoskeletal Systems." In 2023 American Control Conference (ACC), pp. 3005-3011. IEEE, 2023.

[3] Stolpe, Raphael, and Yannick Morel. " Output-prediction Based Nonlinear Control of a Class of Neuro-musculoskeletal Systems." Zenodo (2023): 2023-11.

The work on embodiment that has informed and supported the coherent integration of the above system theoretical results with functional neural models was published in Science Robotics,

[4] Prescott, Tony J., and Stuart P. Wilson. "Understanding brain functional architecture through robotics." Science Robotics 8, no. 78 (2023).

Planning: The developed technology addresses an issue central to practical interactions between robotic system and human worker, the trajectory planning required to support handovers between robot and human. Though the problem has received meaningful attention in the literature, the result in (Iori, et al., 2023) was the first to provide robustness to perturbations (including interruptions and unexpected human movements) while supporting online trajectory generation. This combined ability is expected to prove instrumental to reliable, safe deployment in a practical setting.

<u>Links to peer-reviewed publications or preprints related to the reported outcomes</u>: The work on planning was published in the IEEE Robotics and Automation Letters, and in the International Journal of Social Robotics,

[5] Perovic, Gojko, Francesco Iori, Angela Mazzeo, Marco Controzzi, and Egidio Falotico. "Adaptive Robot-Human Handovers with Preference Learning." IEEE Robotics and Automation Letters (2023).

[6] Iori, Francesco, Gojko Perovic, Francesca Cini, Angela Mazzeo, Egidio Falotico, and Marco Controzzi. "DMP-Based Reactive Robot-to-Human Handover in Perturbed Scenarios." International Journal of Social Robotics 15, no. 2 (2023): 233-248.

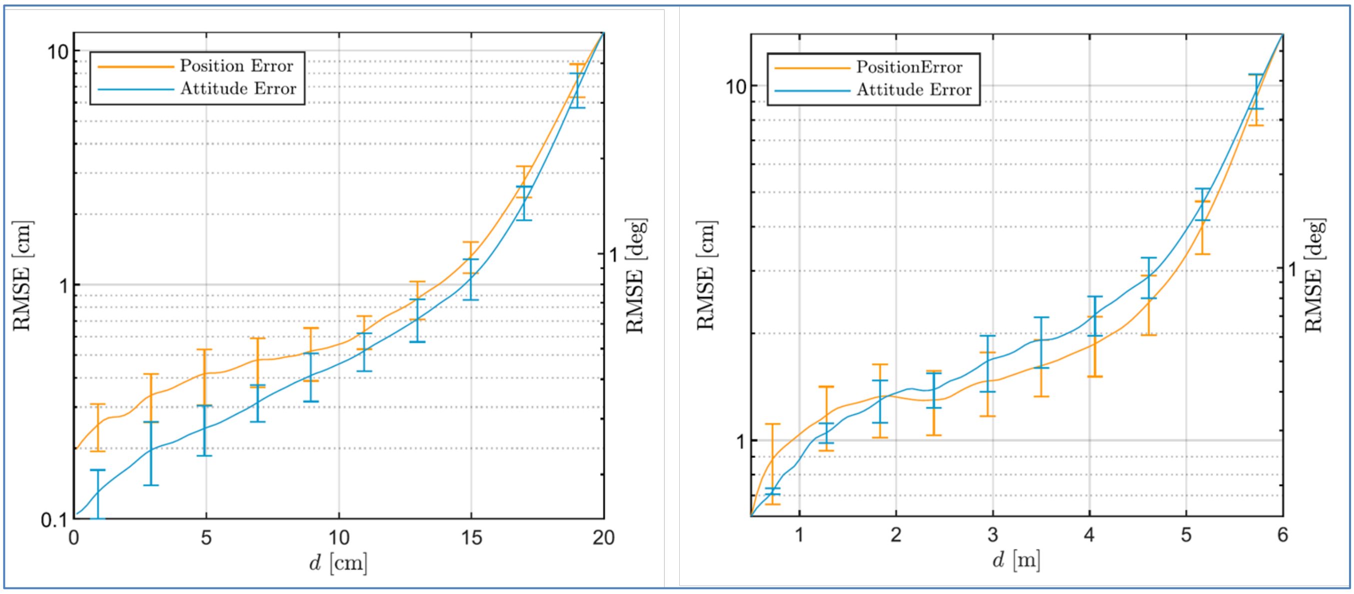


Figure 8: Pose estimation error

Pose estimation error for the haptic (left) and visual (right) modalities implemented in the Showcase demonstrator; the haptic sensors provides centimetric accuracy up to a range of 15cm, whereas ProcNet achieves 2cm positioning error at a distance of 4m (in the absence of occlusion).

Perception: The visual model developed by EPFL (P134) in the first reporting period (ProcNet) achieves robustness to visual occlusion that extends beyond that obtained using alternate existing visual pose estimation models, such as for instance that in (Xiang, Schmidt, Narayanan, & Fox, 2017), on the simulation benchmark we tested it on. Practical relevance hinges on its ability to breach the reality gap, in particular accounting for inter-subject variability in visual aspect of limbs considered. Preliminary experimental findings suggest the impact of such variability on performance remains limited (i.e. centimetric) in controlled, favourable operational conditions. Contribution of the result however also extends to the psychophysics literature, in particular in that our findings confirm some of the insights that supported development of the model, but also infirm others. The implemented capacitive haptic modality, led by UM with active support form USFD, replicates results previously achieved with comparable technology in terms of detection and reflex behaviours (Schlegl, Kröger, Gaschler, Khatib, & Zangl, 2013), with minor contributions in terms of the quality of closed-loop following and avoidance achieved (Zechmair & Morel, Active Electric Perception-based Haptic Modality with Applications to Robotics, 2023). However, the ability to estimate relative pose of objects of interest (see Figure 8, left) is entirely novel and may emerge as a key enabling technology, allowing safe human-robot interactions. Experiments on the physical setup (Figure 5, left) suggest consistent performance across simulation and real world (see illustration on closed-loop behaviours in (Zechmair & Morel, 2023)). The multimodal implementation provides reliable levels of performance, with the combination of ProcNet and haptic modality typically outperforming comparable solutions in situations of stronger occlusion (see Figure 10). Crucially, in situations of sustained, heavy occlusion, information provided from the haptic modality alone is sufficient to guarantee verifiably safe interactions (i.e. allows closed-loop impact avoidance for the entirety of the kinematic relative configuration space explored), which constitutes a meaningful contribution to the robotics literature (manuscript under preparation).

Links to peer-reviewed publications or preprints related to the reported outcomes: The work on haptics was presented at the 2023 IEEE International Conference on Intelligent Robots and Systems, the work on touch from USFD that directly support these results was presented at the Conference on Biomimetic and Biohybrid Systems,

[7] Zechmair, Michael, and Yannick Morel. "Active Electric Perception-based Haptic Modality with Applications to Robotics." In 2023 In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2023.

[8] Salazar, Pablo J., and Tony J. Prescott. "Simple Synthetic Memories of Robotic Touch." In Conference on Biomimetic and Biohybrid Systems, pp. 3-15, Springer Nature Switzerland, 2023.

The work on the extension of deep predictive coding visual models to promote robustness to occlusion is under review for the 2024 IEEE International Conference on Robotics and Automation, a preprint is available on arXiv (doi.org/10.48550/arXiv.2310.18009),

[9] Zechmair, Michael, Alban Bornet, and Yannick Morel. "ProcNet: Deep Predictive Coding Model for Robust-to-occlusion Visual Segmentation and Pose Estimation." arXiv (2023): 2023-11.

Real-Time Execution of SNN for Perception and Motor-Control: in the real-time, physical human-robot co-working demonstrator (KTH (P39)), we implemented a variety of different (spiking) convolutional neuronal networks to simultaneously track a set of three typical workshop objects (here: hammer, screwdriver, pliers). A sufficiently large convolutional SNN for tracking (Figure 9, left) consists of four layers with decreasing representation sizes, followed by a coordinate transformation block to convert a peak of activity into a 2D Cartesian position. Tracking performance of a sample tool in different computing models (ANN, ANN-3, LI neuron, LI-F neuron) is shown as the Cartesian l2 norm (Figure 9, right) between tracked position and ground truth provided by an overhead tracking system.

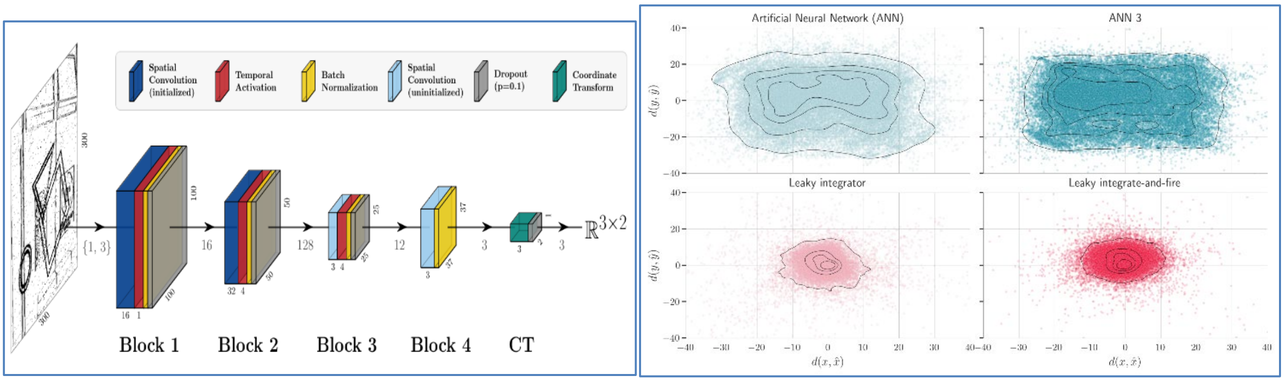


Figure 9: Convolutional Neural Network for object localization (left); localization results (right)

Note that in Figure 9, right, blue colours show ANN performance, whereas red colours show LI (Leaky Integrator) or LIF (Leaky Integrate and Fire) results, respectively. The ANNs operate either on a single 1ms time slice of the event-stream (ANN); or they operate on three consecutive 1ms time slices (ANN-3) to provide a similar signal complexity as neurons with temporal memory (LI or LIF). The resulting plots show a significant better tracking precision of LI/LIF compared to ANN implementations (black contour lines show consecutive 20% margins).

Haptic sensor distance	Camera distance	Level of occlusion	e_{v}	e_{m}	e_{nv}	e_{nm}
Within Range (0-50cm)	Short distance (0–3m)	Light occlusion (0–33%)	1.73cm	**1.59cm**	1.62cm	**1.59cm**
		Medium occlusion (33–66%)	2.48cm	**1.83cm**	4.83cm	2.59cm
		Heavy occlusion (66–100%)	5.84cm	**3.27cm**	6.29cm	3.37cm
	Medium distance (3–6m)	Light occlusion (0–33%)	1.81cm	1.32cm	1.57cm	**1.28cm**
		Medium occlusion (33–66%)	2.59cm	**1.90cm**	5.02cm	2.18cm
		Heavy occlusion (66–100%)	6.04cm	**3.54cm**	6.54cm	4.53cm
	Large distance (6–10m)	Light occlusion (0–33%)	4.07cm	3.06cm	2.65cm	**2.43cm**
		Medium occlusion (33–66%)	5.39cm	4.32cm	2.86cm	**2.90cm**
		Heavy occlusion (66–100%)	–	–	–	–
Out of Range (50cm+)	Short distance (0–3m)	Light occlusion (0–33%)	1.71cm	1.71cm	1.61cm	**1.61cm**
		Medium occlusion (33–66%)	2.45cm	**2.45cm**	4.79cm	**4.79cm**
		Heavy occlusion (66–100%)	5.80cm	**5.80cm**	6.25cm	6.25cm
	Medium distance (3–6m)	Light occlusion (0–33%)	1.83cm	1.83cm	1.56cm	**1.56cm**
		Medium occlusion (33–66%)	2.60cm	**2.60cm**	5.03cm	5.03cm
		Heavy occlusion (66–100%)	5.99cm	**5.99cm**	6.50cm	6.50cm
	Large distance (6–10m)	Light occlusion (0–33%)	3.98cm	3.98cm	2.61cm	**2.61cm**
		Medium occlusion (33–66%)	5.33cm	5.33cm	2.85cm	**2.85cm**
		Heavy occlusion (66–100%)	–	–	–	–

Figure 10: Pose estimation error for visual and multimodal (haptic-visual) perception schemes

With the object located at different distances from haptic sensor and camera, and considering different degrees of occlusion, ProcNet (first column, e_{v}) outperforms PoseCNN (third column, e_{nv}) in situations of medium to heavy occlusion. Complementing vision with haptics, the multimodal solution based on ProcNet (second column, e_{m}) also performs better for medium to heavy occlusion than the PoseCNN counterpart (fourth column, e_{nm}) which perform best in situations of limited occlusion.

Links to peer-reviewed publications or preprints related to the reported outcomes: The work on real-time execution of SNN for perception and motor control led to the following publications,

[10] Turner, James, Jens Pedersen, Jörg Conradt, and Thomas Nowotny. "Event-based dataset for classification and pose estimation." In Proceedings of the 2022 Annual Neuro-Inspired Computational Elements Conference, pp. 101-103. 2022.

[11] Pedersen, Jens Egholm, Raghav Singhal, and Jorg Conradt. "Translation and Scale Invariance for Event-Based Object tracking." In Proceedings of the 2023 Annual Neuro-Inspired Computational Elements Conference, pp. 79-85. 2023.

[12] Pedersen, Jens Egholm, and Jorg Conradt. "AEStream: Accelerated event-based processing with coroutines." In Proceedings of the 2023 Annual Neuro-Inspired Computational Elements Conference, pp. 86-91. 2023.

[13] Romero Bermudez, Juan Pablo, Luis A. Plana, Andrew Rowley, Mikael Hessel, Jens E. Pedersen, Steve Furber, and Jorg Conradt. "A High-Throughput Low-Latency Interface Board for SpiNNaker-in-the-loop Real-Time Systems." In Proceedings of the 2023 International Conference on Neuromorphic Systems, pp. 1-8. 2023.

2.3 Relation to EBRAINS

The work performed contributes to the design and development of EBRAINS in a number of key respects. In particular, it demonstrates, through concrete applicative examples (that is, it *showcases*) the ability of EBRAINS services on *Modelling, Simulation & Computing* to support embodiment (or physical grounding) work of direct relevance to the literature on functional neural modelling, in particular in the area of motor control. In complement, it also provides key examples of meaningful ICT problems being addressed exploiting neuroscience expertise from within the HBP community (representative of the broader community of EBRAINS users). The work conducted helps establish the nature of software tools and services of use in support of such developments. In addition, it provides substantial content in terms of functional models, data sets, frameworks, and closed-loop demonstration examples (all made transparently available to relevant EBRAINS development teams) that may be exploited to populate digital platform services supporting embodiment for neuroscience and application of brain-based technology for ICT. Such content includes: (models) range of functional cerebellar and spinal models developed, demonstration of their combined ability to support motor control for different types of musculoskeletal models, implementations of both versions of trajectory planning models, visual pose estimation models (ProcNet, PoseCNN), trained to estimate pose of two different sets of objects, the digital model of the haptic sensor (in a range of configurations from single sensing surface to ten-by-ten configuration), (framework) the supervised learning framework supporting training of haptic and visual modalities, (demonstrators) **1** demonstration of multimodal haptic-visual pose estimation of human limb (integrating either ProcNet or PoseCNN) illustrating pose estimate behaviour as a function of degree of occlusion (see Figure 12), **2** demonstration of safe-by-design effort exchanges between active-compliant robot and human limb (efforts exchanges estimated using the developed haptic modality, and monitored tendon strains used to constrain robot movements), **3** handover demonstration (using ProcNet for object pose estimate, Figure 11, right, and the robot trajectory planning model developed in the reporting period), and **4** the collaborative assembly of a turbine engine (integrating ProcNet and cerebello-spinal motor control, Figure 11, left).

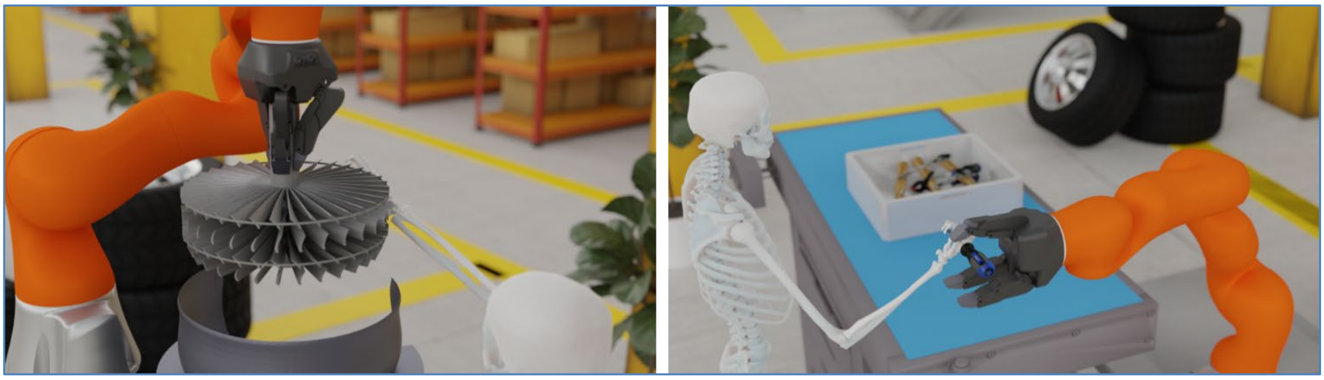


Figure 11: Demonstrations of interaction scenario

Collaborative assembly of turbine engine (left), where the robot provides gravity compensation and guidance towards appropriate assembly position, handover integrating novel planning model (right).

2.4 How to Access the Demonstrator

The simulation demonstrator (including all relevant models and the four demonstrations listed in Section 2.3) can be downloaded from https://drive.ebrains.eu/d/c2ae93eaa8304515b9bf/. Please refer to the document titled *Installation and Execution.odt* (found at the above link) for the requisite steps involved. Videos illustrating the different demonstration scenarii are provided.

3. Perspectives

Activities undertaken in the implementation of Showcase 6 and results obtained have given rise to a wide range of collaborations; several of the addressed areas of investigation provide perspectives of future developments.

Work performed on the implementation of the capacitive haptic modality has found applications in the area of robotic grasping. The technology was adapted to allow integration within robotic manipulators to support shape detection and adaptation (i.e. adjusting digits' configuration to best approach the object to be grasped, promoting emergence of stabler grasps), as well as measure of efforts exchanged (supporting regulation of contact efforts). The resulting gripper system was integrated within a simulation model, together with the digital contact resolution model (Zechmair & Morel, 2022) developed in the previous period (and integrated within the physics of the Showcase demonstrator to support faithful contact physics, see Figure 11). This model was exploited in a reinforcement-learning framework to investigate impact of the aforementioned haptic information on the system's ability to learn stable grasps. The work made use of both classical (*static*) grasp metrics found in the literature, and of the dynamic metric (Zechmair & Morel, 2021) investigated in Task 3.4. Results obtained are shared with CNR and compared with results from their work on automation of learning grasp affordances, reported in D3.17. Areas of synergies, in particular in terms of adjusting the technology in D3.17 to benefit from our findings, have been identified.

Several of the results achieved have motivated discussions with prospective industrial partners with a vested interest in adapting or maturing the technology to address their needs. The grasping technology discussed in the previous paragraph is of direct interest to the manipulation of delicate objects. UM (P117) is in discussion with a prospective industrial partner (from the automation and AI industry) with needs in the area of automation of manipulation in the food industry. In addition, discussions have been engaged on the extension of the perception technology developed to support robust perception and automation in the area of Maintenance, Repair, Overhaul (MRO) and circular economy (industrial partners in: aeronautics, automation). Additional applications were identified in nuclear decommissioning, with UM (P117) in discussion with an interested prospective partner from the nuclear industry.

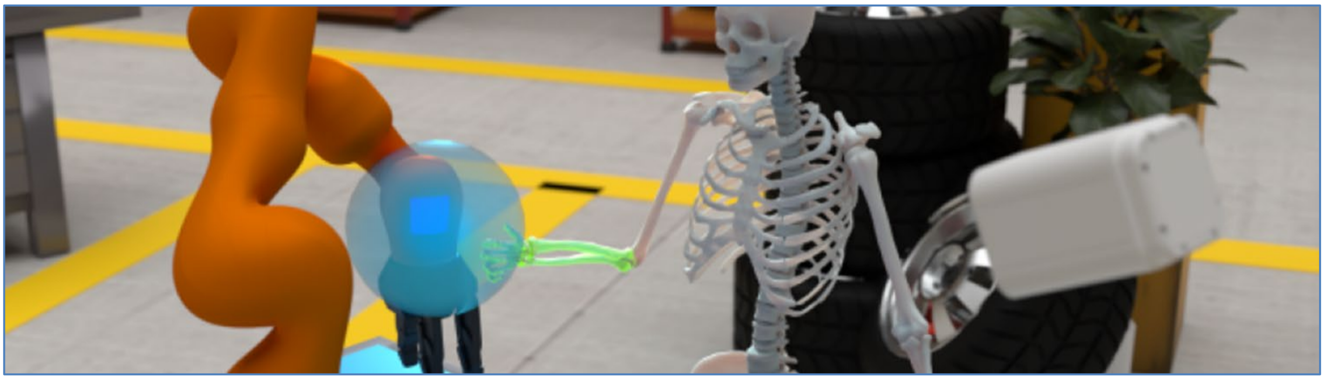


Figure 12: Multimodal pose estimate

Estimated limb pose is represented as a green highlight, visual model used is ProcNet, the haptic sensor's range is represented using a blue translucent body.

Work on functional modelling of neural motor control is ongoing. Activities exploiting the developed system theoretical control framework to investigate merit of spinal pathway models (collaboration between UM (P117) and EPFL (P134)) is providing perspectives of development of improved functional spinal models, afforded verified dynamical motor control features. These extended models, in turn, may give rise to different cerebello-spinal synergies. In addition, integration of the functional (system theoretical) motor control scheme developed in Showcase 6 with additional, complementary levels of representation of relevant neural systems engaged in motor control (i.e. population-level whole-brain model descriptive of evolution of plasticity, and detailed cortical models of loci of plastic adaptation) has been discussed. The intent is for such a multiscale model to be exploited, propagating constraints across scales, to help investigate the relationship from voluntary movements to plastic adaptation in stroke patients.

4. References

Ames, A. D., Coogan, S., Egerstedt, M., Notomista, G., Sreenath, K., & Tabuada, P. (2019). Control barrier functions: Theory and applications. *IEEE European control conference*, (pp. 3420-3431).

Bermudez, J. P., Plana, L. A., Rowley, A., Hessel, M., Pedersen, J. E., Furber, S., & Conradt, J. (2023). A High-Throughput Low-Latency Interface Board for SpiNNaker-in-the-loop Real-Time Systems. *International Conference on Neuromorphic Systems*. (P4057)

Bruel, A., Abadìa, I., Collin, T., Sakr, I., Lorach, H., Luque, N., . . . Ijspeert, A. (2023). The spinal cord facilitates cerebellar upper limb motor learning and control; inputs from neuromusculoskeletal simulation. *bioRxiv*. (P4072)

Cherubini, A., Passama, R., Crosnier, A., Lasnier, A., & Fraisse, P. (2016). Collaborative manufacturing with physical human–robot interaction. *Robotics and Computer-Integrated Manufacturing*.

Haddadin, S., & Croft, E. (2016). Physical human–robot interaction. *Springer handbook of robotics*, 1835-1874.

Hogendoorn, H., & Burkitt, A. N. (2019). Predictive coding with neural transmission delays: a real-time temporal alignment hypothesis. *Eneuro*.

Iori, F., Perovic, G., Cini, F., Mazzeo, A., Falotico, E., & Controzzi, M. (2023). DMP-Based Reactive Robot-to-Human Handover in Perturbed Scenarios. *International Journal of Social Robotics*, 233-248. (P3865)

Knowles, T. C., Stentiford, R., & Pearson, M. J. (2021). WhiskEye: A biomimetic model of multisensory spatial memory based on sensory reconstruction. *Towards Autonomous Robotic Systems*, (pp. 408-418). (P3083)

Kruijne, W., Bohte, S. M., Roelfsema, P. R., & Olivers, C. N. (2021). Flexible working memory through selective gating and attentional tagging. *Neural Computation*. (P3194)

Linsley, D., Kim, J., Veerabadran, V., Windolf, C., & Serre, T. (2018). Learning long-range spatial dependencies with horizontal gated recurrent units. *Advances in neural information processing systems*.

Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. *IEEE conference on computer vision and pattern recognition*, (pp. 3431-3440).

Lorach, H., Galvez, A., Spagnolo, V., Martel, F., Karakas, S., Intering, N., . . . al, e. (2023). Walking naturally after spinal cord injury using a brain–spine interface. *Nature*.

Lotter, W., Kreiman, G., & Cox, D. (2016). Deep predictive coding networks for video prediction and unsupervised learning. *arXiv*.

Luenberger, D. (1971). An introduction to observers. *Transactions on automatic control*, 596-602.

Olson, I. R.-C. (2004). Neuronal representation of occluded objects in the human brain. *Neuropsychologia*, 95-104.

Pearson, M. J., Dora, S., Struckmeier, O., Knowles, T. C., Mitchinson, B., Tiwari, K., . . . Pennartz, C. M. (2021). Multimodal Representation Learning for Place Recognition using Deep Hebbian Predictive Coding. *Frontiers in Robotics and AI*. (P3027)

Pedersen, J. E., & Conradt, J. (2023). AEStream: Accelerated event-based processing with coroutines. *Neuro-Inspired Computational Elements Conference*, (pp. 86-91). (P4030)

Pedersen, J. E., Singhal, R., & Conradt, J. (2023). Translation and Scale Invariance for Event-Based Object tracking. *Neuro-Inspired Computational Elements*, (pp. 79-85). (P4029)

Prescott, T. J., & Wilson, S. P. (2023). Understanding brain functional architecture through robotics. *Science Robotics*. (P4012)

Prescott, T. J., Lepora, N., Mitchinson, B., Pearson, M., Martinez-Hernandez, U., & Grant, R. A. (2020). Active Touch Sensing in Mammals and Robots. *The Senses: A Comprehensive Reference*, 79-109. (P3177)

Rizzotti, A., Jeanneret, L., Studer, B., Bracamonte, J., Ouerhani, N., & Kunze, M. (2023). Learning from demonstration and safe cobotics using digital twins. *Sensors & transducers journal*, 25-32.

Salazar, P. J., & Prescott, T. J. (2023). Simple Synthetic Memories of Robotic Touch. *Conference on Biomimetic and Biohybrid Systems*, (pp. 3-15). (P4171)

Schlegl, T., Kröger, T., Gaschler, A., Khatib, O., & Zangl, H. (2013). Virtual whiskers—Highly responsive robot collision avoidance. *IEEE International Conference on Intelligent Robots and Systems*, (pp. 5373-5379).

Stolpe, R., & Morel, Y. (2023). Model-Based Nonlinear Control of a Class of Musculoskeletal Systems. *American Control Conference* (pp. 3005-3011). IEEE. (P4123)

Strazdas, D., Hintz, J., & Al-Hamadi, A. (2021). Robo-hud: Interaction concept for contactless operation of industrial cobotic systems. *Applied Sciences*.

Thrun, S. (2004). Toward a framework for human-robot interaction. *Human-Computer Interaction*.

Turner, J., Pedersen, J., Conradt, J., & Nowotny, T. (2022). Event-based dataset for classification and pose estimation. *Neuro-Inspired Computational Elements Conference*, (pp. 101-103). (P3772)

Weistroffer, V., Keith, F., Bisiaux, A., Andriot, C., & Lasnier, A. (2022). Using physics-based digital twins and extended reality for the safety and ergonomics evaluation of cobotic workstations. *Frontiers in Virtual Reality*.

Xiang, Y., Schmidt, T., Narayanan, V., & Fox, D. (2017). PoseCNN: A convolutional neural network for 6d object pose estimation in cluttered scenes. *arXiv*.

Zechmair, M., & Morel, Y. (2021). Assessing Grasp Quality using Local Sensitivity Analysis. *IEEE International Conference on Intelligent Robots and Systems*, (pp. 3995-4001). (P4121)

Zechmair, M., & Morel, Y. (2022). Penalty-based numerical representation of rigid body interactions with applications to simulation of robotic grasping. *IEEE International Conference on Electrical, Computer, Communications and Mechatronics Engineering*. (P4122)

Zechmair, M., & Morel, Y. (2023). Active Electric Perception-based Haptic Modality with Applications to Robotics. *IEEE International Conference on Intelligent Robots and Systems*. (P4158)