Showcase 5 - Proof of Concept, performing human-like digit configurations by robotic hand (D3.1 – SGA3)

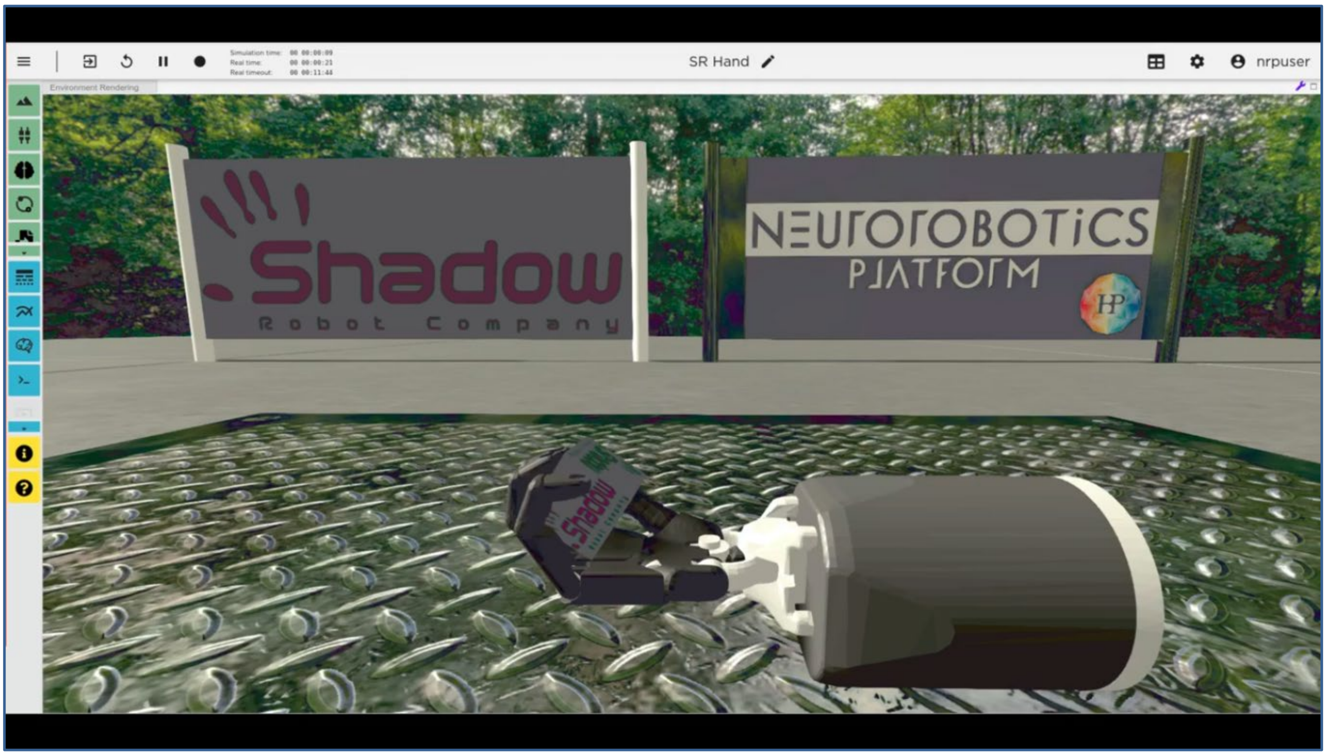

Figure 1: Simulated Shadow Robotic Hand grasping a cube

The Human Brain Project and Shadow collaboration involves integrating Shadow hardware on the HBP Neurorobotics Platform to explore the possibilities offered by neuroscience for motor control, especially with reinforcement learning.

Project Number:	945539	Project Title:	HBP SGA3

Document Title:	Showcase 5 - Proof of Concept, performing human-like digit configurations by robotic hand
Document Filename:	D3.1 (D20) SGA3 M9 ACCEPTED 210504.docx
Deliverable Number:	SGA3 D3.1 (D20)
Deliverable Type:	Demonstrator
Dissemination Level:	PU = Public
Planned Delivery Date:	SGA3 M9 / 31 Dec 2020
Actual Delivery Date:	SGA3 M10 / 19 Jan 2021; resubmitted 16 Apr 2021; accepted 04 May 2021
Author(s):	Mario SENDEN, Yannick MOREL, UM (P117)
Compiled by:	Mario SENDEN, UM (P117)
Contributor(s):	
WP QC Review:	Vaishnavi NARAYANAN, UM (P117)
WP Leader / Deputy Leader Sign Off:	Rainer GOEBEL, UM (P117)
T7.4 QC Review:	Guy WILLIS, Annemieke MICHELS, EPFL (P1)
Description in GA:	Simulation of robotic hand performing human-like digit configurations.
Abstract:	The demonstrator discussed is concerned with dexterity; an anthropomorphic robotic hand performs complex hand movements. A biologically constrained recurrent convolutional neural network (RCNN), providing a stochastic movement policy for each joint of the hand, is specified and trained with reinforcement learning. Biological constraints are considered at a macroscopic and a microscopic scale. At the macroscopic scale, we identify the network of cortical regions involved in the coordination of complex hand movements, their functional specialisation and anatomical connections. We use this information to constrain the number of layers, the computational graph they form and their input modalities. At the microscopic scale, we identify the neuron types and their composition within the involved brain regions. These determine the input-output relations of local circuits and will be used as activation functions of units in different layers of the RCNN. Training of the RCNN occurs *in silico* on high-performance computing clusters.
Keywords:	Sensorimotor architecture, deep learning, motor control, haptics, contact physics, manipulation, robotics.
Target Users/Readers:	General scientific audience with some background in reinforcement learning.

Table of Contents

History of Changes made to this Deliverable (post Submission)

Date	Change Requested / Change Made / Other Action
19.01.2021	Deliverable submitted to EC
04.03.2021	Resubmission with specified changes requested in Review Report Main changes requested: • Change 1: "Before publication, it would be valuable to add a preamble section explaining the positioning of these showcases and related demos in the context of SGA3 and HBP." • Change 2: "In addition, it would be useful for the readers, to better explain what has been done and which cognitive architecture is being pursued and will be demonstrated in this showcase along SGA3 phases. In particular, it should explain why this PoC is a step towards a breakthrough in the near future that may help achieve some of SGA3 external outcomes." • Change3: "Potential risks eventually compromising the completion of the showcase should also be discussed."
31.03.2021	Revised draft sent by SP/CDP to PCO. Main changes made, with indication where each change was made: • Change 1: addressed in Section 1.1 • Change 2: addressed in Section 1.3 & Section 3 • Change 3: addressed in Section 3
16.04.2021	Revised version resubmitted to EC by PCO via SyGMa

1. Context

The work conducted within Work Package 3 (WP3) is structured around a number of integrative demonstrators, which involve contributions from and the active collaboration of most Tasks (and complementary areas of expertise) in the WP. These demonstrators address a range of functions, typically starting from lower abstraction sensorimotor loops, progressing towards cognition. Cognition refers to the ability to attend to external stimuli or internal motivation, to identify the significance of such stimuli, and to execute meaningful responses to them. At the highest levels, this includes hierarchical planning and other functions of comparable abstraction level. Demonstrators in WP3 operate at different levels of description detail and biological plausibility, from detailed spiking models (emphasising biological description) to more abstract rate-based representations (emphasising functional performance). These demonstrators implement a modular approach, helping prototype a supporting modular cognitive framework, in close collaboration with the Service Categories (SCs) of the EBRAINS neuroscience research infrastructure. They emphasize embodiment and consider implementation on Neuromorphic Computing (NMC) hardware where relevant.

1.1 Preamble Showcase 5

The demonstrator (Showcase 5 of WP3) discussed below addresses the more functional end of the spectrum. Specifically, it is concerned with dexterity; an anthropomorphic robotic hand performs complex hand movements such as in-hand object manipulation. To that end, a biologically constrained recurrent convolutional neural network (RCNN), providing a stochastic movement policy for each joint of the hand, is specified and trained with reinforcement learning. Biological constraints are considered at a macroscopic and a microscopic scale. At the macroscopic scale, we identify the network of cortical regions involved in the coordination of complex hand movements, their functional specialisation and anatomical connections. We use this information to constrain the architecture of the RCNN in terms of the number of layers (reflecting the frontoparietal network and the visual system), the computational graph they form and their input modalities (including vision, touch and proprioception). At the microscopic scale, we identify the neuron types and their composition within the involved brain regions. These determine the input-output relations of local circuits (computational units) and will be used as activation functions of units in different layers of the RCNN. Training of the RCNN occurs *in silico* on high-performance computing clusters. The demonstrator is chiefly developed and implemented on EBRAINS in a simulated scenario. This demonstrator provides an end-to-end (deep reinforcement learning) approach towards developing a large-scale, embodied, architecture, subserving cognitive functions in the broad sense as defined above (not extending to higher-level cognition). It thus complements other WP3 activities, which focus on hand-crafting architectures, as well as on individual modules. While this demonstrator is in principle self-contained, we are exploring how it may be expanded upon (with additional layers and functionalities) to serve as a foundation upon which larger architectures requiring dexterity may build. The development of the demonstrator occurs in three stages. First, a proof-of-concept system has been developed to establish the range of required infrastructure services and procedures. Second, a first, working, prototype of in-hand object manipulation is to be trained *in silico*. This stage allows fine-tuning of co-designed infrastructure services and procedures. The final stage involves training the RCNN using domain randomisation of parameters of the hand model and physics simulator, to ensure that the final demonstrator is generalisable towards other physics engines and potentially to physical systems as well.

Situation of Showcase 5 in the SGA3 context: The relation of the demonstrator to the infrastructure is detailed in Section 1.4. Activities directly involved in its development are located in Tasks T3.1 and T3.4. Support on technical aspects is extended by T3.10. Implementation is conducted in collaboration with SC2 (WP4) and SC6 (WP6). Close collaboration with the development team of SC4 (WP5) is ongoing, in particular informing the current Neurorobotics Platform (NRP, SC4) re-factoring to support the type of work involved in the demonstrator's development. The development of the Showcase directly addresses WPO3.1 and PO5. The co-design process with SC4 indirectly supports OC5. More broadly, the workflows and services, to be developed and offered to EBRAINS users as a result of supporting Showcase 5, contribute to OC9. By thus contributing to EBRAINS' unique value proposition, the demonstrator contributes to PO1.

1.2 Specifications of M9 Proof of Concept

Work conducted in the first period of SGA3 focused on setting up the simulation and reinforcement learning framework. To that end, we restricted ourselves to a simplified version of the biologically constrained RCNN (using typical rectified linear activation functions) and considered a simple finger reaching task, wherein the RCNN must join the tip of the hand's thumb with the tip of another randomly chosen finger at a random position above the palm, while all other fingers have to remain in a straight position, approximately on the same plane as the palm.

The stochastic policy instantiated by the RCNN utilises a Beta distribution to provide continuous control. The (digital) model of anthropomorphic hand considered is that of the Shadow Dexterous hand, which represent the technological state of the art for this type of manipulators. The Shadow hand specifies boundaries on the difference in joint angles it can translate in a given period of time. In bounded action spaces such as these, the infinite support of Gaussian-distributed policies biases them towards the limits. The unconstrained support requires a fold of all predictions outside of an action's bounds onto the boundary values and, hence, artificially raises their probability. The Beta distribution's support is limited to the interval [0; 1]. By appropriately scaling samples to the allowed interval of the action, this distribution affords complete and bias-free prediction.

We used proximal policy optimisation (PPO) as reinforcement learning algorithm. Unlike most other implementations of PPO, ours does not embed the algorithm in a general framework, but builds the framework around the algorithm. It is, therefore, lightweight in its API, while going in-depth into optimizations of the original PPO description.

During training, long input sequences to the RCNN can become problematic, because backpropagation through time (BPTT) needs to record the activations throughout the entire sequence in order to calculate gradients. We tackled this problem by using truncated BPTT (TBPTT). This algorithm tackles the issue of long input sequences by dividing each sequence into subparts, over which the gradients are backpropagated. At the end of each subpart, the state(s) of the RCNN state(s) are passed over to the next subpart, but the computation graph is cut off. This alleviates the need for memory but sacrifices accurate modelling of long-distance dependencies. However, for the robotic hand, long-distance dependencies are not of much importance and recurrence is mainly required to encode current environmental conditions and motion directionality.

Training of the RCNN occurs on the EBRAINS high-performance computing infrastructure and uses the MuJoCo physics engine for simulations of the robotic hand. Work is split between workers (utilising CPU nodes) gathering experience and an optimiser (utilising GPU nodes) training the RCNN (policy). A large pool of workers generate experience in parallel by rolling out the current version of the policy in simulations. Workers download the newest policy parameters from the optimiser at the beginning of every epoch, generate training episodes, and send the generated episodes back to the optimiser. The optimiser threads pull down generated experience and stage it to their respective GPU's memory for processing. After computing gradients locally, they are averaged across all threads using MPI, which we then use to update the RCNN parameters.

1.3 Scientific Problems Addressed

The Showcase Demonstrator will lead to breakthrough and scientific outcomes by combining deep learning, robotics and neuroscientific knowledge to understand how the brain coordinates visually-guided object manipulation. Although tool use is common among animals, the significance that tools have acquired for humans is unique. Indeed, skilful object manipulation is an essential component of our everyday life, and our manual dexterity is unmatched by even our closest simian relatives. Given that in-hand object manipulation is a highly complex task that engages a large-scale network encompassing sensory, association and motor regions, it is unsurprising that how the brain achieves this remains insufficiently understood. It has been established that the frontoparietal network performs sensorimotor transformations necessary for goal-oriented action execution. For instance, the intraparietal sulcus (IPS) contains "visual", "visuomotor" and "motor" neurons and sends input to the rostroventral premotor cortex (PMVr). The PMVr stores motor synergies, a low-dimensional vocabulary of motor actions, suggesting that the IPS-PMVr circuit transforms visual information into motor commands. Transformations follow a hierarchical progression, with brain regions storing different intermediate representations. Much can be learned with respect to representations from decoding studies. For instance, recent studies conducted in macaque monkeys using both intracortical electrophysiological recordings and fMRI have shown that object-specific hand configurations can be reliably decoded from monkey homologues of the IPS, PMVr and primary motor cortex. While decoding studies are important to identify representations, they cannot provide information on transformations. However, in order to truly understand visually guided actions, it is pivotal to identify not only representations stored in individual brain regions, but also the computations occurring between brain regions. The central contribution of this Showcase is to shed new light on these transformations by transferring a uniquely human skill to an artificial agent. To that end, we specify an ecologically valid set of visually guided actions and use deep learning to train a recurrent convolutional neural network (RCNN) to perform these actions with an anthropomorphic robotic hand. The architecture of the RCNN is based on functional and anatomical knowledge of the frontoparietal network in the primate brain involved in visually guided hand and arm movements. Training data will be gathered from experience obtained by the agent itself within a simulated environment. The transformations and computations emerging in the trained architecture can subsequently be investigated using tools from neuroscience (e.g. representational similarity analysis), machine learning (e.g. feature visualisation) and data-driven science (e.g. sparse identification of nonlinear dynamics) to provide new insights and hypotheses as to how the brain coordinates complex hand movements.

1.4 Relation to and use of EBRAINS

The involved work directly relies on the HBP's EBRAINS research infrastructure. In particular, it requires high-performance computing (HPC platforms, SC6) for training of the RCNN, as it requires ~100 years of simulated experience and is computationally expensive. Furthermore, design of the RCNN benefits from the human brain atlas (SC2) to identify brain regions involved in complex hand movements, as well as their interconnectivity and local cell profile. This benefits from the development of the "Brainscapes" atlas client, which gives direct programmatic access to data organised within the EBRAINS human brain atlas. It supports the multilevel character of this atlas, which defines cytoarchitectonic maps in multiple reference template spaces at different spatial resolutions (namely the MNI Colin, ICBM152 asymmetric as well as the BigBrain microscopic space), and links them with complementary maps related to brain function, connectivity and fibre architecture.

Conversely, the demonstrator work informs the design of the supporting framework, in terms of specifications. In particular, we continuously communicate with the development team of the NRP (SC4) to provide insights on the MuJoCo physics engine and its relevance for such training tasks, as well as essential requirements for transferring simulation of the RCNN onto the NRP, such as accelerated simulation time and the possibility to programmatically launch several instances of the NRP for deployment on high performance computing clusters.

Crucially, this demonstrator motivates the emergence of infrastructure workflows that span a meaningful range of complementary services, from data services (SC2) to embodiment and function-related services (SC4), with direct support from dedicated compute resources (SC6). No such combination of interoperable services is to be found outside of EBRAINS. Demonstrating the manner in which EBRAINS thus empowers us to break new ground in cognitive computational neuroscience - and its practical application - contributes to establishing the infrastructure's Unique Value Proposition.

By M9, though the presented proof of concept relies on the human brain atlas (SC2) and necessarily requires the dedicated support (and computing resources) of the CSCS HPC cluster (SC6), it makes limited effective use of the NRP (SC4). The development of the Showcase, however, has motivated our active engagement, with the SC4 development team, in the co-design process of the NRP. This is of particular relevance to aspects related to the development of services supporting advanced and brain-inspired AI, allowing to bridge AI, ML, robotics, and neuroscience.

2. How to access the Showcase

The demonstrator is implemented in TensorFlow 2, supported by Python 3.6 and can be accessed from GitHub: https://github.com/ccnmaastricht/dexterous-robot-hand.

A video demonstration can be accessed at https://youtu.be/oPwPOZ1kVC8

3. Perspectives and M21 Specifications

The discussed demonstrator will build upon the M9 Proof of Concept described here in several distinct directions.

1 Biological plausibility of the RCNN: First of all, the biological realism of the RCNN will be improved by replacing activation functions in its layers with input-output relations (firing-current curves) of local circuits within their corresponding cortical regions. This is an important first step to facilitate the transformation of the RCNN into a spiking neural network. A potential risk related to this is that our current understanding of how cell composition of a brain region translates to activation functions is limited, which may slow down the process. To mitigate this risk, the first prototype of the working in-hand object manipulation demonstrator (M21) will be implemented with activation functions analytically derived from identical leaky integrate-and-fire neurons. This work, towards greater biological plausibility, will be supported by exploiting the Brainscapes API, allowing, for instance, exploiting structural connectivity data, afforded by the human brain atlas, to inform design of the RCCN architecture. Similarly, region-specific data will be relied on to derive apposite activation functions. Implementation of the process of constraining the model based on brain atlas data will provide an opportunity to engage partners in the RI work plan to investigate co-design of a more general workflow, promoting the emergence of meaningful, functional connections between neuroscientific data hosted on EBRAINS (in SC2) and functional models (such as those embodied with support from SC4).

2 Analysis supporting transparency and understandability: Related to local dynamics, work in the period between M9 and M21 will also involve identifying governing equations to describe the dynamics of the recurrent layer of the RCNN. To that end, we utilise an approach supporting the sparse identification of nonlinear dynamics (SINDy).

3 End-to-end training for in-hand manipulation: The most important extensions are, however, with respect to the function subserved by the RCNN. Work with regard to these proceeds in two inter-related directions. In the first and primary direction, the RCNN will be trained on in-hand object manipulation. That is, the RCNN needs to learn to manipulate an object (a cube), starting from an initial placement in the hand, to reach a predefined end state (e.g. a specific face of the cube facing towards the camera).

4 Modular approach for in-hand manipulation: In a second direction, the RCNN will be split into two modules. The first module receives tactile, proprioceptive and a desired reach (which finger to join with the thumb) as input and is independently trained on this task. The second module receives visual input and must learn to imitate a visually observed finger tapping sequence by controlling the first module. If the work along the second direction proves fruitful, the approach will also be used to develop a system capable of performing in-hand object manipulation and can thus serve as an alternative to the end-to-end approach that constitutes the core of Showcase 5.

5 NRP integration: Capitalising on the expected development of a MuJoCo API for the upcoming release of the NRP, the functional, trained (either end-to-end or modular approach) network, and its embodiment (physical model of the Shadow Hand) will be made available together in a self-contained NRP model. The eventuality that the NRP MuJoCo integration would be unsuccessful (risk) is mitigated by the extension of the training process to include domain randomisation, allowing to explore generalisability of achieved results, including a range of changes to the physics of the physical agent (which allows considering control of the same robotic hand, simulated using alternate physics engines, from MuJoCo to the NRP-used ODE or Bullet).

6 Integration of vision processing: While the current proof-of-concept demonstrator only included touch and proprioception as input to a recurrent neural network, the M21 demonstrator will include visual processing as well. This will be integrated in the form of a convolutional block pre-trained on an object pose-estimation task. Input to the convolutional block will come from a single camera whose position relative to the hand will be chosen to provide a human-like perspective of the hand and the object it manipulates.

In summary, the M21 demonstrator will feature an RCCN of greater biological plausibility (exploiting SC2 data), whose functions will be extended to support visually-guided in-hand manipulation of a simple object and a functional implementation of the trained functional model will be demonstrated on the NRP.