CDP4 Visuo-Motor Integration - Results for SGA2 Year 2
D2.5.1 - SGA2

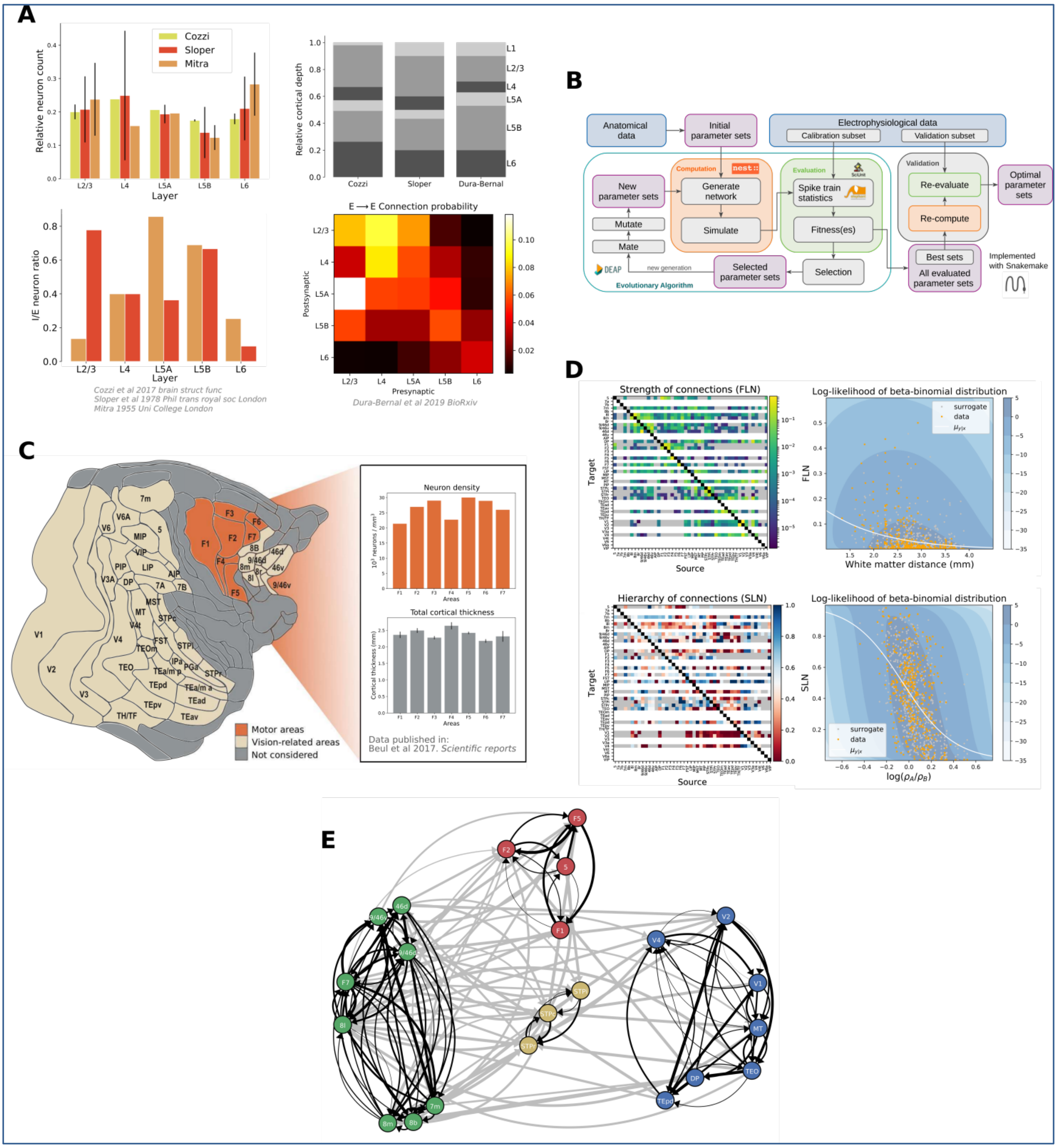

Figure 1: Towards a biologically realistic large-scale model of visuo-motor areas in the macaque.

(A) Primary motor cortex parameters for macaque and mice from the literature [4-8], sample mean values displayed, error bars show standard deviation for the available cases. (B) Workflow for the estimation of the unknown parameters in the model based on electrophysiological recordings. (C) Areas considered in the model (in the M132 parcellation) and characteristics of each motor area. (D) Fractions of labelled neurons (FLN) and supragranular labelled neurons (SLN) [5, 10] for the areas under consideration, and their predictive relation with white matter distance between cortical areas and log-ratios of neuron density, respectively. (E) Community structure of the known connectivity for the areas considered, determined using the map equation method [11]. See KRc4.1, Output 4.

Project Number:	785907	Project Title:	Human Brain Project SGA2

Document Title:	CDP4 Visuo-Motor Integration - Results for SGA2 Year 2
Document Filename:	D2.5.1 (D12.1 D61) SGA2 M24 ACCEPTED 201005.docx
Deliverable Number:	SGA2 D2.5.1 (D12.1, D61)
Deliverable Type:	Report
Work Packages:	WP2.2, WP2.5, WP4.2, WP4.4, WP7.3, WP10.2
Key Result(s):	Krc4.1, Krc4.3, KR2.1
Dissemination Level:	PU = Public
Planned Delivery Date:	SGA2 M24 / 31 Mar 2020
Actual Delivery Date:	SGA2 M24 / 25 MAR 2020; resubmitted 31 Aug 2020; accepted 5 Oct 2020
Author(s):	Mario SENDEN, UM (P117)
Compiled by:	Mario SENDEN, UM (P117)
Contributor(s):	Sacha VAN ALBADA, JUELICH (P20)
SciTechCoord Review:	Martin TELEFONT, EPFL (P1)
Editorial Review:	Annemieke MICHELS, EPFL (P1)
Description in GA:	CDP4 Annual compound deliverable Year 2 For consistent presentation of HBP results, SGA2 M24 Deliverables describing the accomplishments of an entire SP, WP or CDP have been prepared according to a standard template, which focuses on Key Results and the Outputs that contribute to them. Project management elements such as Milestones and Risks will be covered, as per normal practice, in the SGA2 Project Periodic Report.
Abstract:	CDP4 is aimed at understanding the mutual interactions between action and perception with a focus on visual and visually-guided actions such as eye movements and hand-eye coordination. CDP4 is organised around three Key Results splitting it into theoretical (KRc4.1), applied (KRc4.2) and data-driven (KRc4.3) components and combines neuroimaging, neurocomputational modelling, machine learning and robotics. In the current phase, our focus has been largely on theoretical work but progress has been made with respect to the other Key Results as well.
Keywords:	Visuomotor integration, saccades, object recognition, saliency, multi-area modelling, hemispatial neglect, spiking neuron models, deep learning
Target Users/Readers:	Scientists, Companies and other potential users of HBP results. Neuroscientific community, funders

Table of Contents

Table of Figures

History of Changes made to this Deliverable (post Submission)

Date	Change Requested / Change Made / Other Action
24 Mar 2020	Deliverable submitted to EC
30 Jul 2020	Resubmission with specified changes requested in Review Report Main changes requested: • P13. D2.5.1 should be editorially proofread before publication. • P70 Some editorial tuning would be useful before publication of this public deliverable, i.e. avoiding section headings with just numbers (e.g. Output 1).
31 Aug 2020	Revised draft sent by SP/CDP to PCO. Main changes made, with indication where each change was made: • D2.5.1 was proofread • Section headings have been adjusted to be more descriptive
31 Aug 2020	Revised version resubmitted to EC by PCO via SyGMa

1. Overview

Humans are continuously required to interact with a complex, dynamically changing environment. Our brain does thus not operate in a vacuum, but forms a closed loop with the surroundings in which it is embedded. This has important implications for the study of action and perception and their interrelatedness within an ever-changing environment, as one can only be fully understood in context of the others. For example, human vision is only sharp near the centre of fixation and quickly drops with increasing distance from fixation. In order to recognise objects around us and relate their spatial configuration, it thus necessary to perform eye-movements and integrate information over several fixations. A by-product of these eye movements is that static objects move across the retina and the perceptual system needs to correct for object displacement and disentangle it from object motion. Eye movements thus need to be understood in terms of object recognition and vice versa. Co-Design Project 4 (CDP4) is aimed at understanding the mutual interactions between action and perception by combining neuroimaging (using tools such as fMRI), neurocomputational modelling, machine learning and robotics. Our focus lies specifically on visual as well as visually-guided actions, such as eye movements, for object recognition & scene understanding (SGA1 & SGA2) and hand-eye coordination for object manipulation (SGA2: reaching & grasping; SGA3: in-hand manipulations).

With respect to object recognition & scene understanding, we have identified functional components necessary for this task and developed computational models for each of them. Over the past 12 months, we have trained a deep learning architecture to perform scene labelling in light of the aforementioned image blow-up.

With respect to object manipulation, we have set up a recurrent convolutional neural network (RCNN) architecture able to control an anthropomorphic robotic hand as well as a reinforcement learning procedure for training the RCNN to perform in-hand object manipulation.

2. Introduction

The brain enables autonomous agents to interact meaningfully with a dynamic environment. That is, the brain forms a closed loop with the surroundings in which it is embedded through its sensory-motor apparatus. How the integration of sensory and motor function is achieved and how perception and action mutually affect each other constitute important questions in neuroscience. For instance, the sharp drop-off in visual acuity with eccentricity forces the visual system to perform saccadic eye movements and to integrate information across "snapshots" of the visual scene. These eye movements, in turn, affect perception as they lead to blur, retinal displacements and the requirement to distinguish eye- from object movements. Similarly, tasks such as reaching and grasping require tracking of object and hand location in space as well as continuous translation between coordinate frames (e.g. retinotopic vs body-centred).

Co-Design Project 4 (CDP4) fuses computational modelling, deep learning, experimentation and robotics to understand how the brain coordinates such visually-guided actions. To do so, it follows a top-down approach. That is, it starts by identifying and implementing functional components relevant to the task. Implementation may involve developing computational models based on existing neuroscientific data (these modelling efforts occur largely within KRc4.1). It may, however, also involve utilisation of goal-driven deep (reinforcement) learning to let a neural network uncover potential solutions for performing ecologically valid visuomotor tasks (these efforts - largely based on behavioural data and labelled image databases - occur largely within KRc4.3). Subsequently, functional components are integrated into a single large-scale, closed-loop, visuomotor architecture for deployment with robotic systems (KRc4.1). These architectures are continuously refined to increase their biological realism. This occurs in a modular fashion as individual functional components may, for instance, be translated from a rate neuron to a spiking neuron implementation, independent from other components. The architecture may furthermore serve as a virtual patient, to model disorders resulting from damage to the system. The "saccades for object recognition" architecture is especially suited to investigate attention deficits (hemispatial neglect) resulting from stroke (KRc4.2) since it places strong emphasis on attention (saliency) and attention-based decision making (target selection). Co-Design Project 4 was originally organised around three Key Results (KRs): KRc4.1, KRc4.2 & KRc4.3. However, due to unforeseen delays regarding the approval of partnering project BRAINSYNCH-HIT, we, in line with suggestions made by the EC, decided to postpone KRc4.2 until SGA3.

KRc4.1: Visuo-motor integration neuronal network model

By collaborative efforts with SP4 and SP7, efficient network simulations in NEST allow to build novel large-scale neural network models that actually perform challenging visuo-motor integration tasks while being based on state-of-the-art knowledge about the architecture of the brain as well as on novel insights from beyond the state-of-the-art sub-millimetre human functional brain imaging. By collaborative efforts with SP10, the visuo- motor integration model runs on the Neurorobotics Platform. The neurorobotics implementation allows to generate realistic behavioural data.

KRc4.2: Lesioning parietal and frontal areas of eye movement model to explain unilateral spatial neglect stroke and TMS treatment effects

The implemented visuo-motor system will integrate simplified modules of about 20 cortical and sub-cortical areas involved in visual stimulus processing, saliency calculation, target selection and motor planning. The established computational-anatomical relationship and asymmetric attention shifting architecture allows to simulate lesions of the model (link to CDP1) that are related to saliency, attention and motor planning producing neurological symptoms observed in unilateral spatial neglect patients. Furthermore, we will have access to neglect patients that are currently treated with TMS stimulation allowing to model reduced inhibition from the contra-lesional hemisphere. The effects of TMS will be integrated in the neural network model.

KRc4.3: Application of Visuo-Motor Integration Model to User Input Data

The visuo-motor model integrates state-of-the art architectural and functional knowledge of how the brain controls eye movement (version 1 [SGA1]) and grasping (version 2 [SGA2]) supporting the interpretation of the user's neuroimaging, electrophysiological and behavioural data. Since the visuo-motor integration model is implemented in NEST and linked to a (virtual) robotic system, researchers can run the model with their own visual stimuli as input and compare the predicted behaviour with their own empirical data.

3. Key Result KRc4.1 Visuo-motor integration neuronal network model

3.1 Outputs

3.1.1 Overview of Outputs

3.1.1.1 List of Outputs contributing to this KR

1) Spiking neuron model of saccade generator circuit in the brain stem (component title: Visuo-motor integration model performing eye movement and reaching tasks, leader: Rainer GOEBEL, id: C2632, type: model)
2) Integrated "saccades for scene understanding" (SSU) closed-loop architecture (component title: Visuo-motor integration model performing eye movement and reaching tasks, leader: Rainer GOEBEL, id: C2632, type: model, link: https://github.com/ccnmaastricht/CDP4_NRP)
3) Multi-area multi-layer spiking cortical models (component title: Multi-area model of cortical network at neuronal resolution, leader: Sacha VAN ALBADA, id: C730, type: model, link: https://github.com/INM-6/multi-area-model)
4) Recurrent model of V1 orientation tuning in closed loop psychophysics experiment (component title: Visuo-motor integration model performing eye movement and reaching tasks, leader: Rainer GOEBEL, id: C2632, type: model, link: https://github.com/ccnmaastricht/LTI)
5) Visual target selection model (component title: Visuo-motor integration model performing eye movement and reaching tasks, leader: Rainer GOEBEL, id: C2632, type: model, link: https://github.com/ccnmaastricht/target_selection)

3.1.1.2 How Outputs relate to each other and the Key Result

Output 1 and 5 are sub-modules of Output 2: the SSU closed-loop architecture. The latter is a realisation of KRc4.1 as it represents one concrete large-scale neural network architecture that performs a challenging visuo-motor integration task. Output 3 is a key step towards increasing the biological realism of large-scale visuomotor architectures. That is, while Output 2 targets visuo-motor processing from a functional perspective, Output 3 targets it from a biological perspective.

Output 4 constitutes a test case highly useful for developing the Saccades for Scene Understanding closed-loop embodied architecture (Output 2).

3.1.2 Output 1 - Spiking neuron model of saccade generator circuit in the brain stem

During the first year of SGA2, JUELICH started implementing the saccade generation component of the visuo-motor architecture using spiking neurons. This work is now completed.

The model created was inspired by [1] in its overall structure and shares the segmentation into multiple functionally relevant subnetworks with their own dynamical behaviour as well as their interconnections in a biologically plausible fashion. Moreover, the devised method ensures that the model incorporates the single-neuron electrophysiology of the neurons involved in saccade generation (see e.g. [2]) in a phenomenological way and respects Dale's law.

This output has an upstream dependency on component C1857 embedded in Task T4.4.5: Models of sensorimotor integration.

References

[1] Gancarz G, Grossberg S, A neural model of the saccade generator in the reticular formation, *Neural Networks* **11.7**, p 1159-1174 (1998)

[2] Scudder, Charles A., Chris R. Kaneko, and Albert F. Fuchs. "The brainstem burst generator for saccadic eye movements." *Experimental Brain Research* **142.4** 439-462(2002)

3.1.3 Output 2 - Integrated SSU closed-loop architecture

For Task T2.2.1 (Visuo-motor integration model performing eye movement and reaching tasks; component id C2632), researchers at UM, JUELICH, UPF and TUM are currently finalising the integration of all functional modules (ganglion cell image resampling, object recognition, saliency computation, target selection & saccade generation) into an embodied large-scale cognitive architecture able to perform saccades for scene understanding. Once the components communicate successfully with each other as well as the robotic system in which they are embedded, both utilising the 'Integrated Behavioural Architecture' developed by SP10, the architecture will be used to allow a robotic agent in the NRP to provide a description of its environment. We are currently preparing a position paper detailing the top-down, embodied, modelling approach put forth by CDP4 which will include the setup & scene analysis experiments of the SSU architecture as an example.

Manuscript on the SSU is in preparation.

3.1.4 Output 3 - Multi-area multi-layer spiking cortical models

As part of Task T4.2.3 (Multi-area multi-layer spiking cortical models), JUELICH is developing a multi-area model of a cortical network at neuronal resolution (component id: C730) of all vision- and motor-related cortical areas in macaque to study the dynamics of visuo-motor interactions. The model extends the visual multi-area-model [1, 2], where each cortical area is represented by a full-density model of a cortical microcircuit (component id: C944) [3]. Motor areas differ crucially from visual cortex: they have a less prominent layer 4, a far lower neuron density and different internal connectivity. Therefore, it is pivotal to develop a microcircuit of the macaque primary motor cortex. JUELICH has collected anatomical parameters from macaque and mice [4-8] and has identified gaps in knowledge; especially interlaminar connectivity has been imperfectly characterised. Thus, JUELICH is estimating the missing motor cortex parameters from electrophysiological recordings in the macaque (INT Marseille, FORTH Heraklion) using evolutionary optimisation. For the construction of the large-scale visuo-motor model, JUELICH has refined the existing predictive relations between structure and connectivity [5, 9, 10] using a maximum likelihood estimation of a bivariate betabinomial distribution. The improved estimates will be used to predict the missing connectivity parameters and deliver a model that will serve to investigate mechanisms underlying visuo-motor interactions, for instance in SGA3 WP3.2 'Biological visuo-motor architectures on HBP platforms' and in the FLAG-ERA Project 'PrimCorNet' (Characterization and layer-specific modeling of fronto-parietal dynamics in primate cortical networks).

References

[1] Schmidt M, Bakker R, Hilgetag CC, Diesmann M, and van Albada SJ. Multi-scale account of the network structure of macaque visual cortex. *Brain Struct Func*, 223:1409-1435, 2018 (P1036).

[2] Schmidt M, Bakker R, Shen K, Bezgin G, Diesmann M, and van Albada SJ. A multi-scale layer-resolved spiking network model of resting-state dynamics in macaque visual cortical areas. *PLOS Comput Biol*, 14:e1006359, 2018 (P1457).

[3] Potjans TC and Diesmann M. The cell-type specific cortical microcircuit: Relating structure and activity in a full-scale spiking network model. Cereb Cortex 24:785-806, 2014.

[4] Cozzi B, et al. The laminar organization of the motor cortex in monodactylous mammals: a comparative assessment based on horse, chimpanzee, and macaque. *Brain Struct Func* 222:2743-2757, 2017.

[5] Markov NT, Ercsey-Ravasz MM, Ribeiro Gomes AR, Lamy C *et al.* A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb Cortex* 24:17-36, 2014.

[6] Mitra NL. Quantitative Analysis of Cell Types in mammalian neocortex. *PhD Thesis, University College London*, 1955

[7] Sloper JJ, R. Hiorns W and Powell TPS. A qualitative and quantitative electron microscopic study of the neurons in the primate motor and somatic sensory cortices. *Phil Trans Roy Soc London*, 1978

[8] Dura-Bernal S, Neymotin SA, Suter BA, Shepherd GMG, Lytton WW. Long-range inputs and H-current regulate different modes of operation in a multiscale model of mouse M1 microcircuits. *ArXiv*, 2019

[9] Bakker R, Wachtler T, and Diesmann M. CoCoMac 2.0 and the future of tract-tracing databases. *Front Neuroinformatics* 6:30, 2012.

[10] Markov NT, Vezoli J, Chameau P, Falchier A, Quilodran R *et al.* Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. *J Comp Neurol* 522:225-259, 2014.

[11] Rosvall, M., Axelsson, D. & Bergstrom, C. The map equation. *Eur Phys J Spec Top* **178,** 13–23, 2009.

3.1.5 Output 4 - Recurrent model of V1 orientation tuning in closed loop psychophysics experiment

For Task T2.2.1 (Visuo-motor integration model performing eye movement and reaching tasks; component id C2632), researchers at UM developed a model of early visual cortex in closed-loop with an orientation discrimination psychophysics experiment. Specifically, the model is able to judge the orientation of a stimulus with respect to an unseen reference and to learn from feedback received from the experiment. At the same time, the experiment adjusts difficulty (orientation of stimuli with respect to reference) based on the model's performance according to a staircase procedure.

3.1.6 Output 5 - Visual target selection model

Researchers at UM extended a winner-takes-all (WTA) mean-field decision making model to exhibit winnerless competition; i.e to include switching behaviour. We implemented switching by including a negative feedback loop between motor neurons (decision makers) and visual neurons (conveying saliency input). The model is able to generate a sequence of saccade targets for static saliency distributions and can thus account for free viewing behaviour.

3.2 Validation and Impact

3.2.1 Actual and Potential Use of Output(s)

The SSU architecture (Output 2) constitutes one example of large-scale neural network architectures able to perform challenging visuo-motor integration tasks. Its modular nature allows researchers both within (e.g. SSSA in SP10) and outside HBP to extend it by further functional components such as a cerebellar controller as well as to replace existing components with their own models. This allows researchers to evaluate their models in the context of other models together constituting a behaviourally performant encompassing system. The SSU further serves as a starting point for SGA3 WP3.1. Furthermore, Output 3 is highly relevant for SGA3 WP3.2 as well as for the FLAG-ERA Project 'PrimCorNet'. Outputs 1 and 5 are essential components of the Saccades for Scene Understanding closed-loop embodied architecture developed by CDP4 and currently being embedded as a module in that architecture by joint efforts between SP2, SP4 and SP10. Output 4 constitutes a test case for developing closed loops between models and their environments (an experiment). Experience gained during development of this simplified (non-embodied) test case is proving highly useful for developing the Saccades for Scene Understanding closed-loop embodied architecture. Furthermore, researchers at Maastricht University not involved in the HBP are using the model to predict human performance in perceptual learning experiments.

3.2.2 Publications

Lange, G., Senden, M., Radermacher, A., & De Weerd, P. (2020). Interfering with a memory without erasing its trace. *Neural Networks*, *121*, 339–355. https://doi.org/10.1016/j.neunet.2019.09.027 (P2053)

This publication concerns output 4 and confirms that this Output has been validated by scientific peer review.

4. Key Result KRc4.3: Application of Visuo-Motor Integration Model to User Input Data

4.1 Outputs

4.1.1 Overview of Outputs

4.1.1.1 List of Outputs contributing to this KR

1) Object recognition deep learning architecture (component title: Brain-constrained deep learning modules for visuo-motor integration tasks, leader: Rainer GOEBEL, id: C2634, type: model, link: https://github.com/ccnmaastricht/Object_recognition)
2) Improving convolutional neural networks with biologically connectivity profiles (component title: Brain-constrained deep learning modules for visuo-motor integration tasks, leader: Rainer GOEBEL, id: C2634, type: model, link: https://github.com/ccnmaastricht/inhibition-net)
3) Predicting cortical saliency maps from deep encoder-decoder architecture (component title: Brain-constrained deep learning modules for visuo-motor integration tasks, leader: Rainer GOEBEL, id: C2634, type: model, link: https://github.com/alexanderkroner/saliency)
4) Neural network able to learn to translate desirable finger positions to necessary manipulations of the joints of an anthropomorphic hand (component title: Brain-constrained deep learning modules for visuo-motor integration tasks, leader: Rainer GOEBEL, id: C2634, type: model, link: https://github.com/ccnmaastricht/dexterous-robot-hand)

4.1.1.2 How Outputs relate to each other and the Key Result

Output 1 relates mainly to Output 2 of KRc4.1 as it constitutes one functional component of the SSU closed-loop architecture. It further relates to KRc4.3 as it is a deep learning architecture trained based on human behavioural (scene labelling) data. Output 2 bridges the gap between CDP4 and WP3 in SGA3 as it constitutes a first step towards building bio-inspired deep learning architectures which is an important aspect of WP3. Finally, Output 3 is a follow-up of Output 2 of KRc4.3 detailed in SGA2 Deliverable D2.2.1 (deep encoder-decoder architecture for saliency computation). The bio-inspired deep learning architectures (Output 2) have been further developed for the control of finger joints of an anthropomorphic robotic hand (Output 4).

4.1.2 Output 1 - Object recognition deep learning architecture

Researchers at UM have previously developed an algorithm which resamples images in accordance with ganglion cell distributions in the human retina. Resampling leads to image blow-up such that central regions in the image are enlarged while distant regions are compressed. This resampling is incorporated as an initial step in convolutional neural networks to provide them with a human-like visual acuity drop-off with increasing distance from fixation. This forces these networks to explore a visual scene by taking snapshots from different fixations and integrate this information in order to recognise objects. Combining object identity with eye position information as well as working memory of previous objects and their locations using long-short-term memory (LSTM) units allows a recurrent convolutional neural network to not only identify objects but also place them in spatial relation to each other. This architecture was pre-trained on the Flickr 8k (Hodosh *et al*. 2013) data set and is currently being integrated in the SSU architecture where it will be fine-tuned to label scenes in the NRP.

Results of Output 1 form part of the SSU (Output 2 KRc4.1) and will be included in the manuscript currently in preparation on that architecture.

References

Hodosh, Micah, Peter Young, and Julia Hockenmaier. "Framing image description as a ranking task: Data, models and evaluation metrics." Journal of Artificial Intelligence Research 47 (2013): 853-899.

4.1.3 Output 2 - Improving convolutional neural networks with biologically connectivity profiles

Lateral connections play an important role for sensory processing in visual cortex by supporting discriminable neuronal responses even to highly similar features. Researchers at UM have leveraged a biologically inspired Mexican hat lateral connectivity profile along the filter domain to improve the classification performance of convolutional neural networks (including state of the art networks such as the capsule-net). Introducing this connectivity profile has the effect of ordering filters in a sequence resembling the topographic organisation of feature selectivity in early visual cortex.

A manuscript detailing the results of Output 2 has been submitted to the International Conference on Machine Learning (ICML) 2020.

4.1.4 Output 3 - Predicting cortical saliency maps from deep encoder-decoder architecture

Researchers at UM have used the saliency encoder-decoder architecture to generate predictions of cortical saliency distributions in response to natural images are currently conducting a 7 Tesla fMRI study to evaluate these predictions. Specifically, the study aims to contrast saliency maps based on low-level features (such as employed by the Itti & Koch saliency model) and on semantic information (encoder-decoder architecture) on their ability to predict saliency maps observed in posterior parietal cortex and frontal eye fields.

During functional scans participants passively viewed three types of visual stimuli. The first type consisted of grids of simple shapes with the shape in one location differing from the rest with respect to one or two low-level features such as colour, size or shape. The second type consisted of natural scenes with a single salient region (based on eye tracking). The last type consisted of natural scenes for which saliency models based on low-level features and models based on semantic information predict distinct saliency distributions. Pilot scans have been completed and acquisition of the full data had been scheduled for M24. However, due to the COVID-19 pandemic all experimentation involving human participants has been suspended until further notice.

4.1.5 Output 4 - Neural network able to learn to translate desirable finger positions to necessary manipulations of the joints of an anthropomorphic hand

Researchers at UM have developed a biologically inspired recurrent neural network (RNN) for the control of finger joints of an anthropomorphic robotic hand. The RNN must learn to translate desirable finger positions to necessary manipulations of the joints of its hand in order to achieve OpenAI's gym implementation of the reaching task (Plappert *et al*. 2018).

References:

Matthias Plappert, Marcin Andrychowicz, Alex Ray, Bob McGrew, Bowen Baker, Glenn Powell, Jonas Schneider, Josh Tobin, Maciek Chociej, Peter Welinder, Vikash Kumar, and Wojciech Zaremba. Multi-goal reinforcement learning: Challenging robotics environments and request for research, 2018.

4.2 Validation and Impact

4.2.1 Actual and Potential Use of Output(s)

Results of Output 1 form part of the SSU (Output 2 KRc4.1). Output 2 is highly relevant for the machine learning community and all relevant code will be made available. However, because a manuscript detailing this work is currently under double-blind review and code availability would render authors identifiable, code will only be made available after publication of the manuscript.

Once data acquisition and pre-processing is complete, the data of Output 3 will be made publicly available through the Knowledge Graph. Output 4 constitutes a crucial first step towards achieving in-hand object manipulation using the same learning algorithm but a more sophisticated neural network architecture and is thus an important preparation for SGA3.

4.2.2 Publications

Manuscripts submitted or in preparation (see above).

5. Conclusion and Outlook

During the second year of SGA2, all parties involved in CDP4 have made significant progress in Key Result KRc4.1 through the development of a large-scale, embodied, visuomotor architecture able to perform saccades for scene understanding. All functional components are now operational and being integrated in the NRP using the IBA framework. However, fine tuning the scene labelling convolutional neural network for use in the NRP after pre-training on natural images is still a challenge to which we are devoting the majority of our efforts as SGA2 Month 24 approaches. This challenge is exacerbated by the COVID-19 pandemic limiting physical meetings. Nevertheless, we are confident that a first version of the embodied architecture will be achieved by the end of SGA2. The visuomotor architecture has the potential to become an important reference architecture into which other researchers can embed their models of related functions as well as a source of predictions of behavioural data in eye movement experiments.

KRc4.3 also has made good progress. Using biological principles, we were able to improve classification performance of convolutional neural networks. Furthermore, the encoder-decoder saliency architecture continues to be used actively, not only for robotic control but also as a model of the brain as it is used to generate predictions of saliency distributions in response to natural images within posterior parietal cortex and the frontal eye fields measured with 7 Tesla fMRI. Unfortunately, the fMRI experiment had to be suspended due to the COVID-19 pandemic. Lastly, the scene labelling recurrent convolutional neural network is an important component of the SSU and utilises the image distortion algorithm developed throughout the first 12 months of SGA2.

In addition to the aforementioned achievements, the last months of SGA2 have been used to prepare work in SGA3. For instance, as visuomotor work will be continued in WP3, we have begun to develop a biologically inspired recurrent convolutional neural network (RCNN), which will be trained through reinforcement learning to perform in-hand object manipulation using an anthropomorphic robotic hand. The RCNN, as well as the training procedure, are currently being set up on the PizDaint supercomputer at CSCS and work over the next months will be dedicated to troubleshoot and optimise training to ensure a smooth transition to SGA3. Furthermore, the SSU architecture will also be utilised by WP3 in SGA3. Specifically, it will be extended to include a cerebellum module.