Welcome to the final OpenDR Project Newsletter

Newsletter 4.0: Final release of the OpenDR toolkit and other project achievements

Almost everything we hear about artificial intelligence today is thanks to deep learning (DL). Deep learning has achieved tremendous performance jumps in the last decade in several Computer Vision (CV) and Machine Learning (ML) tasks, achieving in many cases super-human performance. However, DL cannot be currently fully exploited in robotics scenarios due to a number of barriers.

Learning Curve Barrier

DL has a **steeper learning curve** than traditional CV and ML methods

Computational Complexity Barrier

DL requires vast amounts of computational power and **energy**

Static Perception Barrier

DL is applied on **static environments** and **does not exploit spatial or temporal embodiment**

The need for an open deep learning toolkit that contains easy to train and deploy real-time, lightweight, Robot Operating System (ROS) compliant deep learning models for robotics is evident. This is where the **OpenDR** project enters.

What is OpenDR?

OpenDR "Open Deep Learning for Robotics Toolkit", is a EU 2020 Project which was launched on January 2020 and aimed to develop a modular, open and non-proprietary toolkit for core robotic functionalities by harnessing deep learning to provide advanced perception and cognition capabilities, meeting in this way the general requirements of robotics applications in the areas of healthcare, agri-food and agile production. The **OpenDR** project was coordinated by the Aristotle University of Thessaloniki, Greece and ran through December 2023 with a total budget of 6.6 Million Euros.

OpenDR aimed to enable real-time robotic visual perception on high-resolution data and enhance the robotic autonomy exploiting lightweight deep learning for deployment on robots and devices with limited computational resources. In addition, it aimed to propose, design, train and deploy models that go beyond static computer vision and towards active robot perception, providing deep human-centric and environment active robot perception, as well as enhanced robot navigation, action and manipulation capabilities.

OpenDR's expected impact is to improve the technical capabilities in robotics by providing easily deployable, efficient and novel Deep Learning tools, as well as to lower the technical barriers by providing a modular and open platform for developing Deep Learning for Robotics tools. Concerning industry, the project's expected impact is to enable a greater range of applications in agri-food, healthcare robotics and agile production, as well as to strengthen the competitiveness of companies by lowering the cost to access robotics-oriented Deep Learning tools.

OpenDR Consortium

OpenDR consortium was a very good mix of 8 partners from 7 European Countries: 2 companies working in various fields of robotics, one company working in the field of robotics simulations, and 5 Universities that join the project with 4 robotics laboratories and 3 deep learning and computer/robot vision laboratories.

Aristotle University of Thessaloniki (AUTH) is the largest university in Greece, established in 1925. AUTH coordinated the project and led the organization of dissemination activities. AUTH focused its research on deep human centric active perception and cognition, where it contributed on deep person/face/body part active detection/recognition and pose estimation, deep person/face/body part tracking, human activity recognition, as well as social signal analysis and recognition. AUTH also led the research in object detection/recognition and semantic scene segmentation and contributed to other areas such as evaluation and benchmarking activities of the project.

Tampere University (TAU) is Finland's second-largest university with 20.000 students and 330 professors. TAU participated with two labs/groups namely the Laboratory of Signal Processing at the Department of Computing Sciences and the Cognitive Robotics Group at the Department of Automation Technology and Mechanical Engineering. TAU led the research in deep human centric active perception and cognition, working mainly on deep speech and biosignals analysis and recognition, and contributed to deep person/face/body part active detection/recognition and multi-modal human centric perception and cognition as well as in a number of other topics. TAU also contributed on defining the agile production use case requirements and specifications and on the integration of OpenDR to this use case.

University of Freiburg (ALU-FR) is one of Germany's leading research institutions with an international reputation in many fields. ALU-FR led the research in deep environment active perception and cognition. ALU-FR focused its research on Deep SLAM and 3D scene reconstruction, as well as on deep navigation. It also contributed on developing methodologies for deep planning.

Aarhus University (AU), Denmark participated in OpenDR with two groups, namely the Data-Driven Analytics Group and the Artificial Intelligence in Robotics Group, both belonging to the Section of Electrical and Computer Engineering. AU led work on 2D/3D Object localization and tracking and worked on sensor information fusion, as well as object detection/recognition and semantic scene segmentation and understanding. AU also contributed to a number of areas such as deep person/face/body part active detection/recognition, deep person/face/body part tracking, deep planning, etc.

Delft University of Technology (TUD) is the oldest and largest technical university in the Netherlands. TUD led/organized the research activities on deep action and control, deep planning, as well as deep navigation. Furthermore, TUD led and undertook the research activities on human robot interaction. Finally, it led and organized the toolkit evaluation and benchmarking activities of the project.

Cyberbotics (CYB) is a Swiss spin-off company from EPFL, which has been developing the Webots robot simulator since 1998. CYB led efforts of defining the toolkit's requirements and specifications. CYB also worked on developing simulation environments and collecting data. Finally, it also led toolkit integration by collecting and integrating all the OpenDR modules developed by the partners.

PAL Robotics (PAL) is a Spanish SME that provides robotic products and services. PAL organized and coordinated the toolkit integration, as well as the use cases integration activities. PAL also contributed on defining the healthcare robotics use case requirements and specifications and worked on the integration of OpenDR Toolkit to this use case, as well as on its evaluation.

Agro Intelligence APS (AGI), Denmark organized and coordinated the toolkit evaluation, as well as the use cases specific toolkit evaluation activities. AGI also contributed on defining the agri-food use case requirements and specifications and worked on the integration and evaluation of OpenDR Toolkit in this specific use case.

Work Performed in the 4th Year

Toolkit Integration & Evaluation

Following the first release of the toolkit on M24, the OpenDR consortium released five more public versions of the toolkit until M48 (six public versions in total), through its GitHub repository. The latest versions focused on expanding the number of available methods, as well as providing performance improvements to existing ones. In addition, several other improvements were made, including the addition of modular installation options, the support for newer CUDA versions, the implementation of a refined ROS/ROS2 interface, etc. The final version of the OpenDR toolkit was released on December 2023, includes several new tools and provides new functionalities, such as gesture recognition, FSeq2 non-maximum suppression, continual SLAM, intent recognition, object detection class filtering, object detection models for agricultural use cases, RL-based active perception, adaptive high-resolution pose estimation, and others. Furthermore, several existing tools have been enhanced by fixing bugs and including additional demos and ROS nodes. Also, the installation process is now more robust and efficient, while the development pipelines have been significantly improved allowing for more easily developing new tools. The OpenDR team will continue supporting the tool in the years to come. We do believe that this major outcome of the project has fulfilled its goal of providing easy to train and deploy, real-time, lightweight, Robot Operating System (ROS) compliant deep learning models for robotics. Indeed, reception of the toolkit from the robotics / deep learning /computer vision community is already very encouraging: so far, the GitHub repository was awarded more than 550 stars, was forked 86 times and the toolkit (as a whole or individual tools) has been downloaded more than 17000 times since its first release in December 2021.

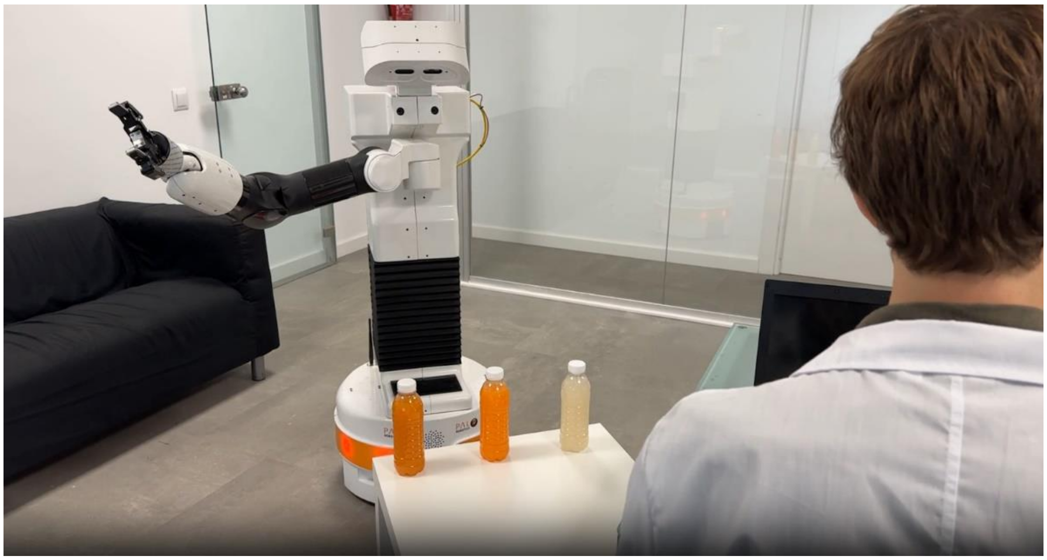

Figure 1: Human robot interaction in the healthcare scenario.

Moreover, the consortium successfully concluded the integration of tools in the three use cases. In the Healthcare use case, specifications were refined to enhance human-robot interaction and multi-modal information fusion, accompanied by a more intricate storyline. In addition, PAL seamlessly incorporated new tools into the healthcare scenario, achieving milestones like dynamic face recognition and improved speech recognition. The integration fostered a seamless interaction between natural speech recognition and PAL's internal chatbot which has resulted in a user-friendly interface for human-robot interaction with the TIAGo robot.

Furthermore, AGI has integrated and tested several tools, meeting the requirements for the Agriculture use case. The Crop & Weed tool was integrated into the OpenDR toolkit and into AGI's Robotti and so did tools for detecting people and tractors and the plant row guidance system. Finally, the Agile Production use case saw updates integrating multimodal tasks, refining simulation-to-real requirements, and introducing natural language recognition.

In terms of evaluation, the consortium rigorously tested and evaluated in diverse environments the performance of a multitude of tools developed in the research work packages. At the same time, tools used in the three use cases were evaluated to check whether they fulfil the respective requirements.

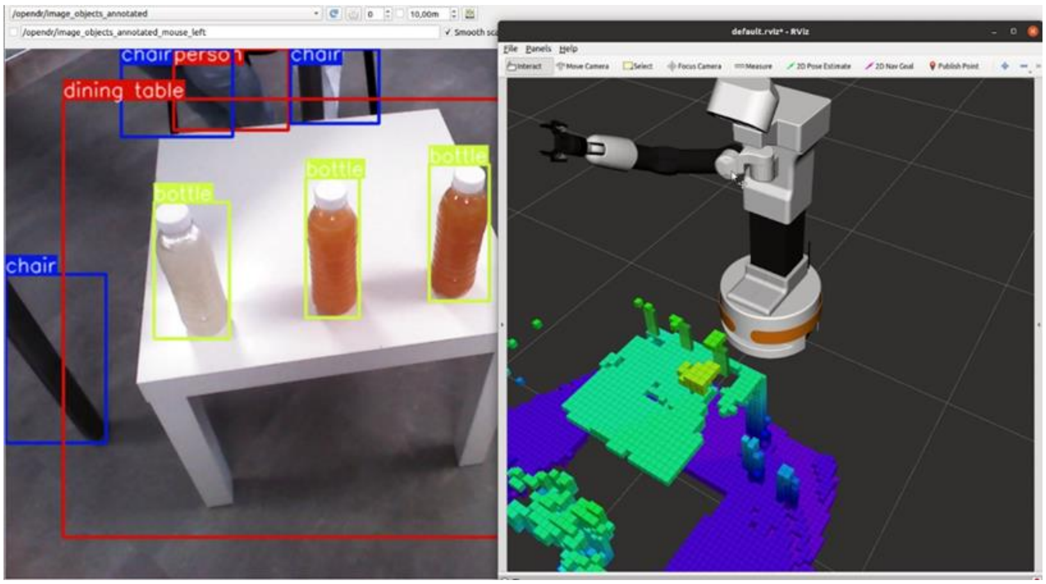

Figure 2: Yolov5 OpenDR 2D object detector (left) used in PAL Advanced grasping pipeline.

As a result, feedback and crucial insights were provided to the partners that developed them. Overall, the integration and testing of the tools, both in the toolkit, and in the three use cases was very successful.

Deep Human Centric Active Perception and Cognition

Work on human-centric active perception and cognition during this last period led to significant achievements. For example, AU worked on incorporating uncertainty estimation in skeleton-based human action recognition by proposing variational

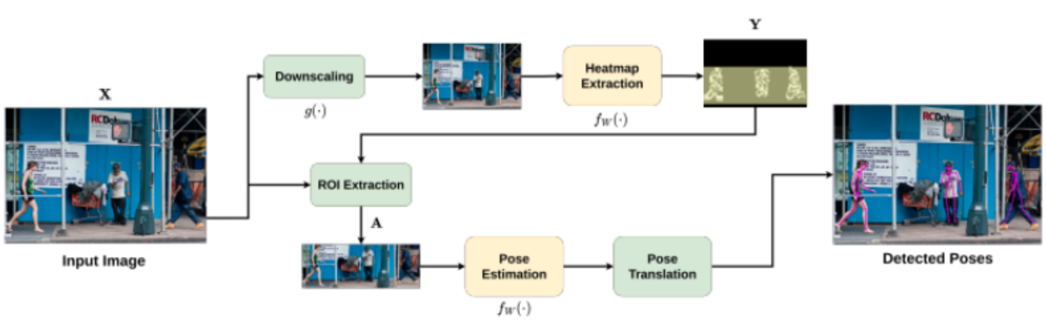

Figure 3: The active perception for high resolution pose estimation approach by AUTH

versions of Spatio-Temporal Graph Convolutional Networks and Adaptive Graph Convolutional Network models.

The proposed variational perception methods improve the quality of human action recognition and estimate model uncertainties. AUTH worked towards further extending a high resolution pose estimation approach developed in the project in order to more efficiently handle cases where multiple humans appear. Furthermore, AUTH further developed an embedding-based active perception approach for face recognition by leveraging a new dataset it has generated. The improved method allows for multi-axes control of the robot.

Finally, TAU developed a method to allow for better performance in a variety of unimodal deep learning tasks (such as hand gesture recognition, audiovisual emotion recognition, language sentiment analysis, etc.) by training the models in a multimodal fashion.

Deep Environment Active Perception and Cognition

In the final year of the project, partners AU, AUTH, ALU-FR, TAU, and AGI proposed various deep learning-based methods for environment active perception and cognition. In detail, AU introduced multiple methods for estimating and exploiting uncertainty in 3D perception tasks, e.g., by utilizing variational neural networks.

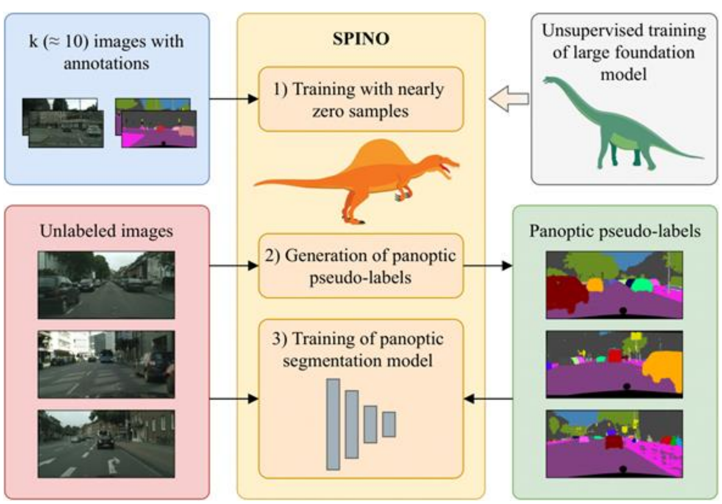

Figure 4: SPINO is a novel method for vision-based panoptic segmentation that enables few-shot learning by exploiting rich image features from an unsupervised foundation model.

Their work significantly improved the tracking of objects. Furthermore, AUTH developed lightweight object detection algorithms and models for agricultural applications while focusing on high-resolution processing. Moreover, AUTH continued working on methods to tackle one-hot encoding shortcomings by introducing a framework for learning soft label embeddings. Partner

ALU-FR proposed several methods for applying online continual learning to visual odometry, monocular depth estimation, and panoptic segmentation. These approaches enable seamless domain adaptation and are self-improving during the test time. ALU-FR further developed a novel 3D scene graph for urban driving that is collectively assembled by multiple agents and provides a high-level structure of large environments.

Additionally, ALU-FR proposed a novel approach for few-shot panoptic segmentation that requires as few as ten annotated images for training. Their method, SPINO, leverages rich image descriptions from a visual foundation model for label-efficient training. Next, TAU prepared benchmark evaluations for a multi-modal feature fusion framework that addresses recent data collection issues in both real and simulated domains. Finally, AGI developed and evaluated a new method for the mapping of agricultural fields and utilized this map for the autonomous navigation of a robot. Using in-house tools, the crop rows are automatically found along with the cross-track error of a row.

Deep Robot Action and Decision Making

In the field of deep robot action and decision making, partners AU, ALU-FR, TUD and AUTH continued and finalized their work towards the design of novel, state of the art navigation, planning, and control algorithms. AU presented a novel Deep Reinforcement Learning (DRL) training strategy for addressing robot navigation problems by leveraging the principles of the Lyapunov theory and a novel approach for learning the low-level flight control of UAV robots. ALU-FR introduced an approach that learns to coordinate and combine previous OpenDR works on exploration and mobile manipulation to autonomously solve a novel interactive multi-object search task. Moreover, AUTH developed a DRL-based end-to-end trainable agent for differential-drive wheeled robot navigation, while also developing the appropriate techniques to improve learning efficiency. Furthermore, AUTH finalized

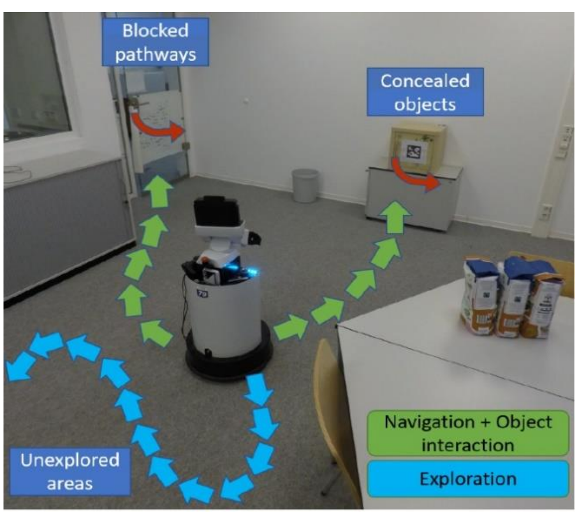

Figure 5: TUD Evaluation of PARTNR in a table-top pick and place task

their work on a data-efficient DRL approach for robust inertial-based UAV localization. TUD focused on mitigating the sim2real gap and enhancing sample efficiency in off-policy reinforcement learning. In addition, TUD worked on simplifying and making safer the collaboration between humans and robots in an industrial context, whereas TAU's main objective was to utilize perception for human-robot collaboration, either by individual perception tools or by combining input from multiple tools into a fused output. Human speech, gestures and object perception provided the input for commanding robot actions and enabling collaboration in shared industrial tasks.

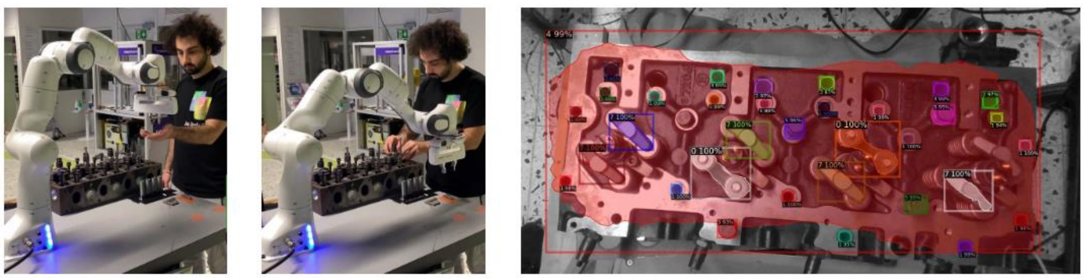

Figure 6: Results of robot-human hand-over and assembly tracking experiments proposed by TAU. (a) hand-over of a rocker arm from robot to human. (b) human assembly action of the rocker arm, while the robot fetches another rocker arm. (c) assembly tracking results, with several objects and their locations detected inside the Diesel engine bounding box.

Simulation Environments and Data

During this year, several models for simulating a robotics farming environment were developed by AU. These include models of robots: AGI's Robotti, tractor implements, agricultural fields, agricultural buildings, farm animals, humans, etc. as well as scenarios based mainly on the use of lidar and camera sensors. The animals, humans, buildings, fences, etc. are used mainly as obstacles that can be perceived by the robot sensors and avoided or causing the robot to stop to avoid a collision. Such an environment can be used to generate large volumes of simulation data that could be processed on-line (during the simulation) or off-line (handy for deep learning purposes and repeatability.

Moreover, AUTH dealt with the creation of a synthetic multimodal (audio, video and 3D models) dataset suitable for active facial expression recognition. In addition, AUTH worked towards a tool that uses Augmented Reality (AR) technology and allows the easy creation of mixed (real and synthetic) data depicting realistic 3D human models in various real environments, captured from different camera positions.

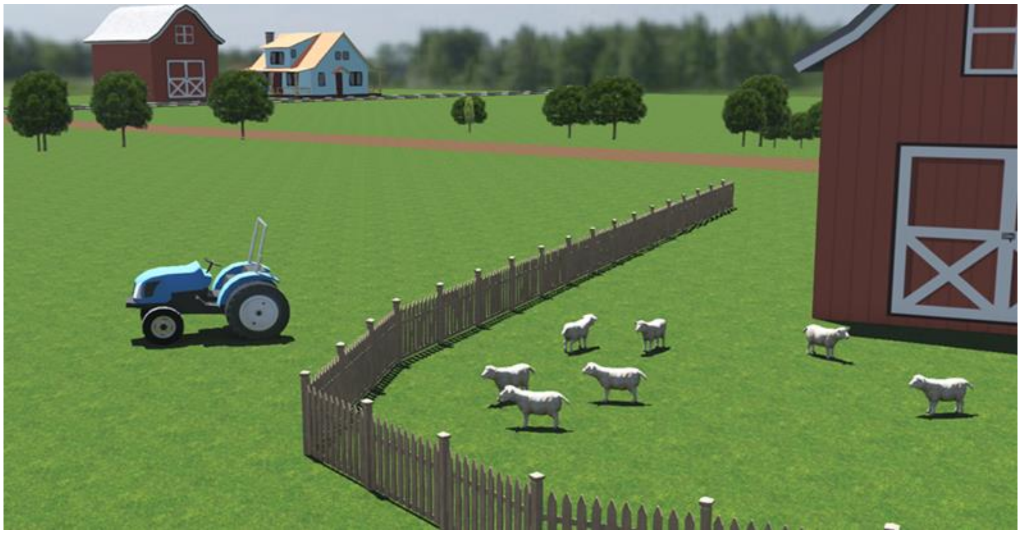

Figure 7: Simulated farm environment, including tractors, animals, barns, etc

AUTH also finalized its work on the Unity-based synthetic annotated data generation pipeline for active vision tasks and publicly released both the software and two datasets (a face datasets called ActiveFace and a full body dataset called ActiveHuman).

Figure 8: Robotti robot operating on a field and detecting the presence of humans.

Finally, CYB worked on further developing the Webots simulator and a new version was released: R2023b. The improvements include the adjunction of new primitive nodes, the integration of the TIAgo robot from PAL in Webots including new device models such as the Robotiq grippers 2f 85, 2f 140 etc. Moreover new versions of the Webots ROS 2 interface with several improvements were released whereas the

Webots web interface was improved with the ability to set up competitions running in webots.cloud.

Dissemination

During the last year of the project, a multitude of efforts, in various directions, were undertaken by the consortium in order to attract interest in the project findings and results. The project website and its social media accounts (Facebook, Twitter, LinkedIn, YouTube) continued to receive frequent updates and posts regarding new publications, project news, forthcoming events etc. Virtually all project publications but also videos and slides from conference presentations are available on the website.

Moreover, the consortium succeeded in generating, for yet another year, a high volume of publications. Eighteen papers were presented or accepted in highly esteemed international conferences (IROS, ICIP, EUSIPCO etc.) and eight papers were published or accepted in scientific journals (including IEEE T-RO and IEEE RA-L). In total, project partners have published 99 conference and journal papers, one edited book and numerous preprints. Not bad at all!

A particularly important dissemination activity was the organization, in collaboration with the MARVEL H2020 research project, of the 4-day "Deep Learning for Autonomous Systems and Smart Cities" Summer School, in Aarhus, Denmark, which was attended by more than 35 persons and included a hands-on introduction to the OpenDR toolkit. Other significant dissemination activities included 2 invited/plenary talks and participation, through talks or posters in partners' booths, in events like ERF 2023, IROS 2023 and ROSCon23.

Epilogue

A very productive four-year journey has come to an end! The project achieved, with hard work and effective partners collaboration, its goal of creating a modular, open and non-proprietary toolkit for core robotic functionalities by harnessing deep learning to provide advanced perception and cognition capabilities. The very good reception of the toolkit (more than 550 stars and 17000 downloads so far) shows that the toolkit has already had a significant impact in the robotics and deep learning communities. Apart from this, the consortium believes that the scientific outcomes of OpenDR, documented in 100 publications, have pushed the state of the art in various areas such as deep environment and human-centric active perception and cognition, deep robot action and decision making, simulation environments and data generation etc.

It was a very interesting journey indeed! We hope that our contributions will be a small but important step towards smarter and more efficient robots of all kinds.

opendr.eu	@OpenDR	OpenDR H2020 Research Project	opendr-eu/opendr	OpenDR

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No *871449*.