Welcome to the 3rd OpenDR Project Newsletter !

Newsletter 3.0: Toolkit progress, integration and evaluation and current status

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 aims 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 is coordinated by the Aristotle University of Thessaloniki, Greece and will be running throughout December 2023 with a total budget of 6.6 Million Euros.

OpenDR will 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 aims 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 is 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 coordinates the project and leads the organization of dissemination activities. AUTH focuses its research on deep human centric active perception and cognition, where it contributes 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 leads the research in object detection/recognition and semantic scene segmentation and contributes 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 participates 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 leads the research in deep human centric active perception and cognition, working mainly on deep speech and biosignals analysis and recognition, and contributes 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 contributes 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 leads the research in deep environment active perception and cognition. ALU-FR focuses its research on Deep SLAM and 3D scene reconstruction, as well as on deep navigation. It also contributea on developing methodologies for deep planning.

Aarhus University (AU), Denmark participates 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 leads work on 2D/3D Object localization and tracking and works on sensor information fusion, as well as object detection/recognition and semantic scene segmentation and understanding. AU also contributes 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 leads/organizes the research activities on deep action and control, deep planning, as well as deep navigation. Furthermore, TUD leads and undertakes the research activities on human robot interaction. Finally, it leads and organizes 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 leads efforts of defining the toolkit's requirements and specifications. CYB also works on developing simulation environments and collecting data. Finally, it also leads 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 organizes and coordinates the toolkit integration, as well as the use cases integration activities. PAL also contributes on defining the healthcare robotics use case requirements and specifications and works on the integration of OpenDR Toolkit to this use case, as well as on its evaluation.

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

Work Performed in the 3rd Year

Toolkit Progress, Integration & Evaluation

The second version of the OpenDR toolkit has been officially released at the end of December 2022. **Version 2.0** of toolkit is more modular, allowing the users to install only the tools they need and includes several new tools like fall detection, SiamRPN 2D tracking, Continual Transformer Encoders, to cite some of them. Additionally, a ROS/ROS2 interface has been added for all tools and the documentation was enriched with instructions and benchmarks. The toolkit is accessible in **GitHub**, as well as through **pip** and **Docker Hub**. Reception of the toolkit from the robotics / deep learning /computer vision community was very good: so far, the GitHub repository was awarded more than 300 stars from its users, whereas the toolkit as a whole or individual tools have been downloaded more than 7500 times since its first release in December 2021. The OpenDR consortium is eager to receive your feedback, bug reports and suggestions for improvements!

Regarding integration, partners AUTH and PAL worked on the integration of the OpenDR toolkit in the different supported embedded platforms (NVIDIA Jetson TX2, AGX, and NX) focusing on the testing and integration of the toolkit in docker for all the platforms to make it easy to use and install. PAL also focused on testing the ROS nodes and provided feedback to partners involved in tools creation towards improving them and making them user-friendly for ROS users. Furthermore, all partners developed and contributed ROS2 nodes, making the OpenDR toolkit ROS2-compliant and extending the range of supported configurations. AGI focused on evaluating the performance of the tools related to the agriculture use case and gathered developer's feedback for improving the interface and documentation of the toolkit.

Deep Human Centric Active Perception and Cognition

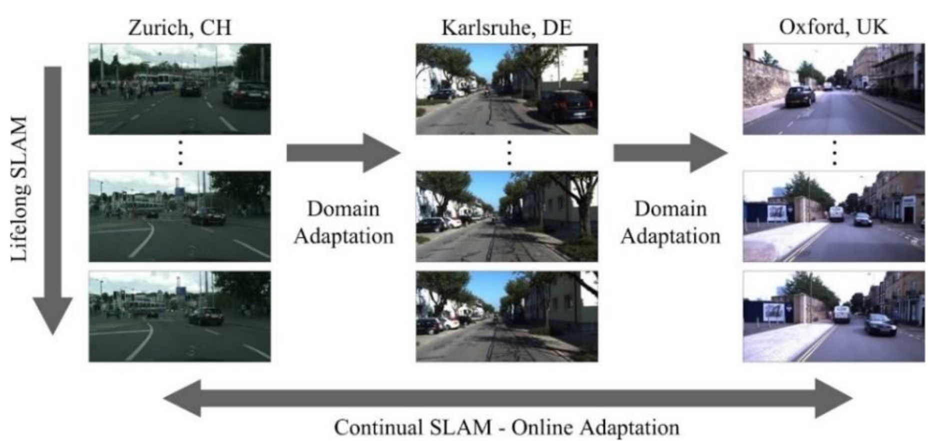
Throughout the third year of the project partners AUTH, AU and TAU have continued their work on human centric tools and algorithms, bringing the consortium even closer to the realization of the powerful, flexible and efficient robotics toolkit. Indeed, AUTH worked towards developing active perception models using a multitude of methods, including active perception for face recognition, high resolution pose estimation, as well non-maximum suppression suitable for person detection methods.

Heatmaps generated as part of the active efficient high resolution pose estimation approach proposed by AUTH.

The developed approaches include a DRL-based control approach for training agents that are able to identify and focus on task-relevant objects, i.e., humans, as well as issue appropriate control commands so as to acquire better results. AU has contributed novel methods for accelerating the online inference of both video- and skeleton-based human activity recognition networks, and also proposed and implemented Continual Inference Networks (CINs), which perform efficient step-by-step online processing. In the area of facial expression recognition, AU researchers proposed a mechanism for improving the generalization ability of the state-of-the-art models on unseen samples by learning diversified facial feature representations and encouraging the learner to extract diverse spatial and channel-wise features. Finally, in multimodal perception, TAU has contributed a method for robust audiovisual emotion recognition. The method includes a new modality fusion approach based on self-attention. Additionally, a training approach, referred to as modality dropout, that increases robustness to incomplete or noisy data in one modality was introduced.

Deep Environment Active Perception and Cognition

During 2022, partners AU, AUTH, ALU-FR, TUD, AGI and TAU proposed a variety of methods that can be used for environment active perception and cognition. AU worked on designing approaches for deep learning models uncertainty estimation, introducing the Layer Ensembles method. AUTH developed a novel deep active object detection pipeline that provides active perception capabilities to existing object detectors by employing a separate planning network that regresses the rotation and translation that a robot should follow in order to increase the object detection confidence. Moreover, AU and AUTH joined forces to introduce a 3D single object tracking method called Voxel Pseudo Image Tracking (VPIT) that is suitable for real-time object tracking using embedded devices. Partner ALU-FR introduced a novel method for 3D multi-object tracking using a graph neural network called Batch3DMOT, which leverages multiple modalities including camera, LiDAR, and radar, and provides 3D tracks of multiple objects in a given scene, as well as a deep learning-based loop closure detection and point cloud registration approach for LiDAR-based SLAM systems. They also dealt with a new task called continual SLAM that combines lifelong SLAM with online domain adaptation to reflect challenges that occur when deploying a SLAM system to the open world and proposed a dual network architecture called CL-SLAM for this task.



Continual SLAM: A new task that extends both lifelong SLAM and domain adaptation techniques.

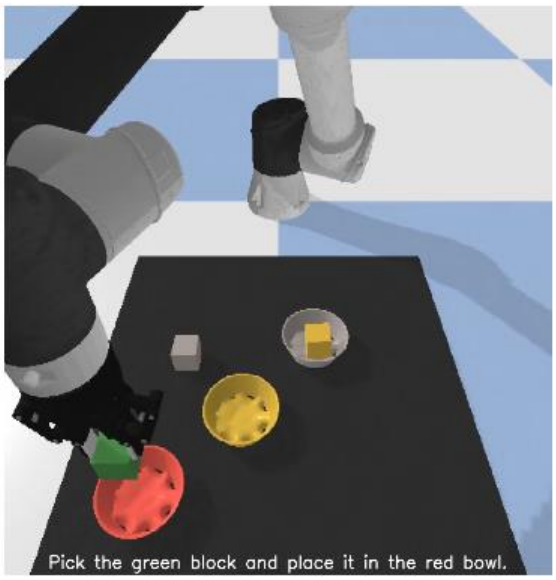
Finally, TUD extended its previously proposed RSH data augmentation scheme for object detection in harsh lighting conditions towards polygon-shaped masked data augmentation in order to further improve robustness against a larger variety of lighting perturbations.

Deep Robot Action and Decision Making

In the area of deep robot action and decision making, partners AU, ALU-FR, TUD, TAU continued to work towards the design of novel navigation, planning, and control algorithms, thus contributing to the state of the art and to the key project objectives. AU presented an end-to-end planner trained with DRL for safe navigation in cluttered obstacle environments. The end-to-end planning algorithm is trained and tested in comprehensive simulations developed in Webots. The method was also successfully deployed in real-world indoor environments successfully and the corresponding experiments demonstrated that the proposed UAV planner trained solely with simulation can directly work in a real environment. On another front, ALU-FR developed a novel multi-object search approach that unifies short- and long-term reasoning in a single model. In addition, ALU-FR has developed an active localization method that combines differentiable particle filters with reinforcement learning to scale to large maps and continuous action spaces.

ALU-FR real world experiments on the PR2 (left) and HSR (right) robots.

TUD further improved its EAGERx toolkit, part of the OpenDR toolkit. The framework has now a consistent interface with an interactive GUI, unit tests with code coverage > 95%, and is accompanied by extensive documentation including a set of 10 interactive tutorials to make it easy for new users to get started. Moreover, TUD presented a delay simulation framework that allows delays to be accurately simulated in simulators that run faster than real- time. Furthermore, a novel communication protocol was proposed that reduces the effect of jitter on the sim2real performance. Finally, TUD started working on the PARTNR human-robot interaction algorithm, which learns to solve ambiguities in pick and place problems through interactive learning.

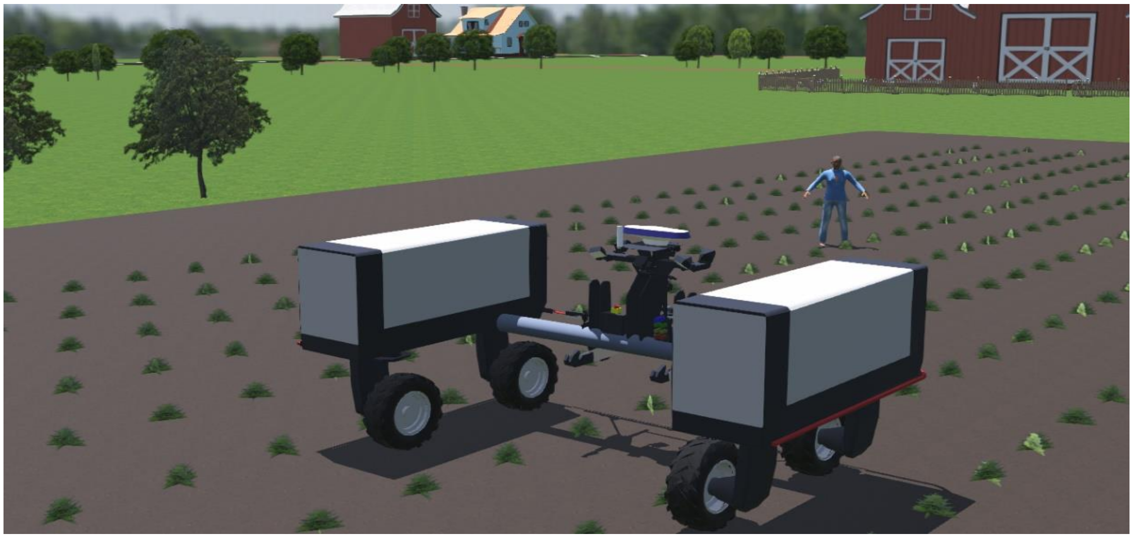


TUD Evaluation of PARTNR in a table-top pick and place task

TAU has made improvements to the SingleDemoGrasp tool. The tool now includes data annotation and augmentation functionalities as part of the toolkit, and the functionality to utilize different visual detection modules from the Detection 2D tool. Furthermore, TAU has developed a human-robot collaborative scenario that is inspired by the agile production use case. Within this scenario, perception tools (human skeleton detection, human action recognition detection & pose estimation of objects and targets) from the OpenDR toolkit are utilized to enable the robot to act as assistant to the human, by functionalities such as automated pick-and-place and robot to human hand-overs. TAU additionally provided basic HRC software templates, which can be utilized to replicate their experiments.

Simulation Environments and Data

Within the previous year, CYB extended the features of the Webots simulator by adding four new robots, several new assets to model industrial, healthcare and rural environments. The Camera node now supports either spherical or cylindrical projections for more realistic simulations and a new CadShape node that allows to easily import 3D models in Collada (.dae) or Wavefront (.obj) format was introduced.



Robotti on a field

Improvements were also made in the way <extern> controllers communicate with the simulator allowing controllers to run over TCP or in isolated docker containers, decoupling the need to run everything on the same machine. Furthermore, a new service https://webots.cloud was announced which allows to easily disseminate results in the form of immersive 3D animations or simulations without needing to install any software. Furthermore, AUTH continued working towards the generation of datasets and simulation environments related to OpenDR use-cases and tools. This work included the generation of 3D facial models, a simulation environment and an image sequences dataset suitable for active facial expression recognition methods, an annotated synthetic full-body /face dataset for human-centric active vision tasks, and a video data generation framework for UAV detection methods. In addition AUTH, along with AGI, created a dataset designed for the training and evaluation of the performance of person detection methods in settings encountered in the agriculture use case. Finally, AU continued to improve the capability of its planning algorithms in its Webots R2021a implemented pipeline by incorporating OpenDR perception tools (2D object detection algorithms). As a result, the safety of algorithms increased by detecting the human presence and stopping the operation in such case.

Dissemination

Numerous efforts, in various directions, were undertaken by the OpenDR 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.

The consortium managed to generate, for yet another year, a high volume of publications. Eighteen papers were presented or accepted in well-established international conferences (including ECCV, ICIP, ICPR, EUSIPCO etc.) and six papers were published or accepted in scientific journals (including IEEE RA-L and Neurocomputing). In total, project partners have published so far 74 conference and journal papers, one edited book and numerous preprints.

A particularly important dissemination activity was a tutorial on "Open and Trustworthy Deep Learning for Robotics" that was organized by the project at IROS 2022, Kyoto, Japan. The tutorial was well attended and included talks on open Deep Learning tools for robot perception navigation and control, as well as other topics, presented mainly from consortium members. Of equal importance was the organization, in collaboration with the FOCETA H2020 research project, of the 5-day "Continuous Engineering and Deep Learning for Trustworthy Autonomous Systems" Summer School, which was attended by more than 75 persons, Other significant dissemination activities included invited talks, talks in two tutorials (at ICRA and CCTA 2022) and participation, through workshop talks or posters / videos in partners' booths, in events like ERF 2022, IROS 2022 or ICRA 2022.

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*.