AI for Industry at the VANDAL Lab of Politecnico di Torino

Giuseppe Averta[1,2], Raffaello Camoriano[1], Barbara Caputo[1,2], Carlo Masone[1,2], Tatiana Tommasi[1,*], Fabio Cermelli[2,†], Enrico Civitelli[3,†], Viviana D'Alto[4,†], Riccardo Ercoli[5,†], Emmanuele Lacavalla[6,†], Daniele Pillan[7,†], Giorgio Pochettino[8,†] and Eugenio Realini[9,†]

[1]Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

[2]Focoos AI, Corso Castelfidardo 22, 10128 Torino, Italy

[3]COMAU, Via Rivalta, 30, 10095 Grugliasco (TO), Italy

[4]STMicroelectronics, Via Tolomeo 1, 20007 Cornaredo (MI), Italy

[5]Aditus, Via Caboto 35, 10129 Torino, Italy

[6]Intesa Sanpaolo, Corso Inghilterra, 3 10138 Torino, Italy

[7]CMA Robotics, Viale del Lavoro 41 Z.I.U, 33050 Pavia di Udine, Italy

[8]Italdesign, Via Achille Grandi 25, 10024 Moncalieri (TO), Italy

[9]GRED, Via Cavour 2, 22074 Lomazzo (CO), Italy

Abstract

We present the research activities that the Visual and Multimodal Applied Learning (VANDAL) Laboratory of Politecnico di Torino is performing in collaboration with several national and international companies. The partnerships foster innovation in areas that range from deep learning model efficiency to robotic manipulation, autonomous driving, and remote sensing.

Keywords

computer vision, multi-cue integration, machine learning, intelligent systems, robotic manipulation

1. Introduction

Politecnico di Torino hosts different research groups working on Artificial Intelligence (AI). Their work is supported by European and National projects, as well as by many local and international enterprises interested in including AI solutions in their production processes. In this contribution, we present the *Visual and Multimodal Applied Learning* (VANDAL) Laboratory at the Department of Control and Computer Engineering and highlight ongoing industrial collaborations.

The research team at VANDAL focuses on the development of the body of theoretical knowledge and algorithms that elaborate on vision and multi-modal signals to support the decision-making process of intelligent agents. We expect these agents to understand the nuanced real world and be able to safely interact with it thanks to robust models that generalize across data variability, can deal with data scarcity, manage novelty, and incrementally learn new concepts. Overall these learning models should also efficiently adjust to the available computational resources, from edge devices to exascale. The developed learning approaches are applied to a large number of tasks such as semantic segmentation for autonomous driving, vision-based geolocalization from ground and satellite images, anomaly detection for product defects, egocentric action recognition, 3D object design and generation. Besides models and methods for perception, VADAL also develops innovative approaches for robotic action considering advanced reinforcement learning algorithms that guide rigid and soft end-effectors for object manipulation and task-oriented grasping.

VANDAL members lead the Turin Unit of the *European Laboratory of Learning and Intelligent Systems*[1] (ELLIS) network and the Artificial Intelligence Hub of Politecnico di Torino [2]. Moreover, they are also involved in the didactic activities of the National PhD in Artificial Intelligence, in particular within the path of Industry 4.0 which is led by Politecnico di Torino.

2. Industrial Collaborations

2.1. FocoosAI

The VANDAL Lab has a deep interest in technology transfer of research activity toward relevant applications for the market. In particular, a successful experience lies in the contribution to the tech development of FocoosAI[3], a spin-off company of Politecnico di Torino, co-founded by faculties of the lab and former PhD candidates. FocoosAI

✉ tatiana.tommasi@polito.it (T. Tommasi)

[1]https://ellis.eu/
[2]https://ai-h.polito.it
[3]https://www.focoos.ai

aims at revolutionizing the development of efficient yet effective neural architectures for complex computer vision tasks, such as image segmentation on single and multimodal data. Moving from the research carried out by the VANDAL lab in the last few years, FocoosAI is developing a proprietary automatic neural design framework that optimizes the cost/accuracy trade-off of deep models considering end-users' requirements in the process. Such technology, for which a patent is currently pending and filled by our lab members, showcases superior performances in the optimizations of neural networks at large, demonstrating the feasibility of deploying complex computer vision tasks even on very low-power devices [1]. Thanks to funding coming from VCs, FocoosAI is currently developing an automatic platform to serve our IP as a b2b service, aiming at proposing a service for a variety of integration and end-user developers that will use FocoosAI products to optimize their AI-based vision pipelines for efficient deployment.

2.2. STMicroelectronics

STMicroelectronics[4] is a global leader in semiconductor design and integrated device manufacturer, with 14 main manufacturing sites worldwide, focusing on the development of energy-efficient products and solutions to boost the intelligence of everyday products.

ST technologies enable smarter mobility, more efficient power and energy management, and the wide-scale deployment of the Internet of Things and connectivity. ST is among the major industrial players pushing toward the deployment of solutions on the very tiny edge, showcasing several use cases of machine learning algorithms on ultra-low power devices for predictive maintenance, computer vision, multi-modal processing etc.

The Vandal Lab shares the interest for the development of efficient deep learning solutions, and a fruitful collaboration between the lab and the company is investigating the feasibility to deploy complex computer vision tasks on ultra-low power, HW accelerated devices. We focused on image semantic segmentation, a popular and industry-relevant CV task where the model is asked to assign a class to each pixel of the image. The collaboration is studying the sensitivity of mainstream semantic segmentation models to network quantization and pruning, targeting the development of standardized compression techniques for existing models specifically tailored for very tiny edge deployment such as ST microcontrollers and microprocessors with dedicated Neural Processing Unit.

2.3. Intesa Sanpaolo

Intesa Sanpaolo (ISP[5]) is one of the top banking groups in Europe, with a significant Environmental, Social and Governance (ESG) commitment, a world-class position in Social Impact, and a strong focus on climate. AI naturally provides relevant support to the elaboration of the massive amount of data managed by the bank and in these processes it is crucial to keep both model fairness and efficiency under control. VANDAL is currently collaborating with ISP for the development of models to evaluate the environmental impact of distributed learning algorithms. In this setting, the learning process takes place on a network of devices with limited resources, and their information is combined into a single model. Distribution helps to reduce some of the data maintenance costs and avoids privacy issues compared to centralized approaches. Additionally, by optimizing communications and operational conditions, it is possible to achieve a significant reduction of the energy footprint, overcoming the need for centralized infrastructures such as cooling or power supply. The collaboration is supported by a Cascade Innovation Grant of the Project *National Centre for HPC, Big Data and Quantum Computing* within the National Recovery and Resilience Plan (PNRR) and is performed in collaboration with other laboratories of Politecnico di Torino (Prof. Michela Meo, Department of Electronics and Telecommunications).

2.4. EFORT

EFORT Europe[6] is the operational holding company that handles business development for the overseas market of the international group EFORT Intelligent Equipment Co. Ltd., a leader in robotics and industrial automation. Within the EFORT group, CMA ROBOTICS[7] designs, produces, and installs a wide range of robotic painting systems that integrate human-teaching programming: the computer control records the operator's movements and repeats them at the programmed speed. Still, these approaches require a dedicated learning phase at each new object instance. Existing approaches to reduce human supervision are either based on simplified premises on object shapes or require expensive offline optimization routines, which hinder their practical applicability for industrial production lines. The collaboration with VANDAL started from the need to overcome these limitations and led to the development of a novel 3D leaning-based framework capable of dealing with arbitrary 3D surfaces and handling a variable number of unordered output spray-painting paths. Our approach predicts local path segments, which are later concatenated to reconstruct

---

[4]https://www.st.com/content/st_com/en.html

[5]www.intesasanpaolo.com/
[6]https://efort-europe.com/
[7]www.cmarobot.it

long-horizon paths [2]. We showed experimentally that the predicted output paths cover up to 95% of previously unseen object surfaces, even without explicitly optimizing for paint coverage. The ongoing collaboration targets a patent for the designed learning model.

2.5. Comau

Robotic manipulation presents significant challenges, involving the completion of intricate perceptual and action-based subtasks. Existing learning approaches rely on a high amount of supervision and do not generalize across different scenarios. Indeed, conventional robot learning methods train a separate model for every application, every robot, and even every environment. Moreover, models tend to be computationally expensive, especially when involving multiple perceptual modalities that are essential to characterize the observed scene and actions. Comau[8] is a leading company in the industrial automation field at a global level and develops Industry 4.0-enabled systems, products, and services. In collaboration with VANDAL, Comau aims to develop a versatile multi-task grasping system for object pick and place. Depending on the observed scene and reliable grasping points, the proposed learning model will select the most suitable end-effector to be used by a multi-gripper robotic arm. Moreover, new high-precision methods for pose estimation from 3D point clouds and 2D images are being investigated to improve the reliability of robot interaction with target objects with cost-effective hardware and sensors.

2.6. Italdesign

Italdesign[9] is a renowned provider of services dedicated to the development of new vehicles and industrial products. The company has rich expertise at all levels of the vehicle development process, ranging from the conceptualization and design to the assembly and construction, and also development and integration of electrical/electronic/software components and passive/active safety systems. In particular, Italdesign has a resident team dedicated to assistive and autonomous driving solutions for which it is interested in exploiting cutting-edge AI solutions. This has prompted a collaboration with the VANDAL Lab, which has established expertise in developing data-driven perception solutions, with a focus on robustness and reliability. The collaboration has led first to an extensive assessment of semantic segmentation methods in driving applications under varying deployment conditions (e.g., in terms of lighting or weather conditions) [3], and then to the development of a new algorithm for data-efficient adaptation from synthetic to real scenarios [4].

---

[8]www.comau.com
[9]www.italdesign.it

2.7. Geomatics Research & Development

GReD[10] is an SME that studies, designs, and develops innovative and highly customized solutions based on geomatic technologies. In particular, GReD developed an innovative system for geodetic and environmental monitoring of structures, ground, and natural hazards, mainly based on an accurate GNSS positioning (Global Navigation Satellite System). It allows to automatically and continuously compute the position of points on the structures or territory with an accuracy of 1 mm on a daily basis and a few mm on hourly basis. Still, only a few points can be monitored, thus the collected metrics are sparser than what is provided by satellite InSAR (Interferometry with Synthetic Aperture Radar) data. On the other hand, InSAR faces constraints in temporal continuity, due to the orbital schedules of satellites. The collaboration between the VANDAL lab and GReD seeks to develop an AI-driven monitoring system that combines InSAR with GNSS measurements. By harnessing both technologies' strengths, our goal is to enhance the detection and forecast of Earth surface deformations and to refine the assessment of landslide risks. The research is performed within the Project EASTERN (EArth obSer-vation models for weaThER eveNt mitigation) supported by a Cascade Grant of the innovation ecosystem *Nord Ovest Digitale E Sostenibile* (NODES) funded by the Ministry of University and Research (MUR) as part of the National Recovery and Resilience Plan (PNRR). The activities also involve collaborations with other laboratories of Politecnico di Torino (Prof. Lia Morra, GRAINS[11]) and the technology consulting company aizoOn[12].

2.8. Aditus and Pro-Logic Informatica

Aditus[13] is an SME that has developed a platform for building personalized cultural experiences (from the online purchase of a ticket using a simple, secure and assisted procedure to the use of a wide range of services). Aditus' platform is set up to promote "Brand Italy" and its enormous potential: from the promotion of Museums and Archaeological Parks to the creation of events, the digitalisation of information and marketing. Aditus and VANDAL are collaborating within the perimeter of the project MAPP (Musei Alpini Phygitali e Partecipativi) (NODES) funded by the Ministry of University and Research (MUR) as part of the National Recovery and Resilience Plan (PNRR). The project aims to develop an innovative model of museum experience based on a physical, digital and participatory approach in two use cases at the Gran Paradiso National Park and the Protected Ar-

---

[10]www.g-red.eu/
[11]https://grains.polito.it/
[12]www.aizoongroup.com
[13]https://aditusculture.com/en

eas Management Body of the Cottian Alps. The project intends to develop a digital ecosystem at the service of outdoor tourism based on the location on the territory of a series of low-cost transit counters capable of providing information on tourist flows to the park authority and at the same time communicating with the tourist in transit via an app/webapp rich in educational digital content and environmental protection, and the use of Visual Place Recognition. VANDAL contributes to the project with the development of AI algorithms for visual place recognition. The activities also involve collaborations with other laboratories of Politecnico di Torino (Prof. Luigi De Russis, e-Lite[14]) and the technology consulting company Pro-Logic Informatica[15], an SME that provides software solutions and support in the design of technological infrastructures. Pro Logic produces hardware and software solutions for the management of data on the Internet (IoT), both for industry and domotic applications, for environmental monitoring (Smart City), and for monitoring industrial processes (Smart Factory).

Acknowledgments

References

[1] N. Cavagnero, G. Rosi, C. Cuttano, F. Pistilli, M. Ciccone, G. Averta, F. Cermelli, Pem: Prototype-based efficient maskformer for image segmentation, in: IEEE/CVF Computer Vision and Pattern Recognition Conference (CVPR), 2024.

[2] G. Tiboni, R. Camoriano, T. Tommasi, Paintnet: Unstructured multi-path learning from 3d point clouds for robotic spray painting, in: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2023.

[3] E. Alberti, A. Tavera, C. Masone, B. Caputo, Idda: A large-scale multi-domain dataset for autonomous driving, IEEE Robotics and Automation Letters 5 (2020) 5526–5533.

[4] A. Tavera, F. Cermelli, C. Masone, B. Caputo, Pixel-by-pixel cross-domain alignment for few-shot semantic segmentation, in: IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2022.