PhD in Computer and Control Engineering

Research Title: Fast and robust visual semantic segmentation for autonomous driving applications

Funded by	ItalDesign

Supervisor	Barbara Caputo barbara.caputo@polito.it Giorgio Pochettino giorgio.pochettino@italdesign.it

Contact	https://scholar.google.it/citations?hl=en&user=mHbdIAwAAAAJ&view_op=list_works

Context of the research activity	Semantic segmentation has recently become one of the most prominent tasks in computer vision. Indeed the ability to assign a label to each pixel of an input image is crucial whenever a very detailed description of the observed scene is needed, as in autonomous driving. Although convnets have obtained significant results, current methods still lack the ability to run at real time (or even below such threshold) on devices equipped with low resources computational devices. Moreover, the ability to segment semantically images regardless of the visual domain in which they are acquired is still below what is needed for real commercial exploitation, as current methods suffer considerably from changes due to varying weather and illumination conditions. This PhD project will attack these issues, aiming to significantly move beyond the current state of the art on these two issues. We will leverage over recent advances in machine learning, fast semantic segmentation and domain generalization to develop approaches able to reach a strong degree of generalization while at the same time keeping the computational load compatible with the needs of

	deployment on a running autonomous agent. Experiments will be conducted on publicly available benchmark databases, as well as on a running prototype.

Objectives	The objectives of this PhD are twofold: (a) Develop a visual semantic segmentation deep architecture able to perform robustly with respect to wide variations in imaging conditions due to varying weather and illumination conditions, and (b) Make such architecture usable in real time (or below) scenarios, as it is needed to use these algorithms onboard of autonomous systems like autonomous cars. To achieve the first objective, we will leverage on the literature in visual domain adaptation and generalization, casting the problem of semantic segmentation as classification at the pixel level across different visual domain. We will first move from the current state of the art of single source domain adaptive semantic segmentation to multi-source scenarios, as this has been shown to increase robustness and generality in domain adaptation. We will then move to the more challenging domain generalization scenario, where the system cannot access during training time any data from the scene where the system will be eventually deployed. In both settings, we will leverage over very recent work proposing self-supervised learning approaches for achieving adaptation and generalization [1]. We will further push the concept for allowing the algorithm to learn to generalize at test time, taking advantage of the ongoing experience of the agent. The result will be a new generation of algorithms for semantic segmentation, able to adapt and generalize to new visual domains. To achieve the second objective, we will tap into recent advances in fast semantic segmentation [2] and principle reduction of network parameters [3] and incorporate these findings into the architecture developed in the first objective. Here the challenge will be to compact the structure of the deep network without losing in accuracy and generalization abilities.

	All work will be tested on publicly available databases as well as implemented and tested on a real prototype. To this end, a large effort for the creation of a new database for semantic segmentation, capturing a wide variety of visual conditions, is currently under way with the close collaboration of ItalDesign. Likewise, the company will guide and supervise the implementation of various versions of the algorithms developed in this thesis on a proprietary prototype for road testing. [1] F. M. Carlucci, A. D'Innocente, S. Bucci, B. Caputo, T. Tommasi. Domain generalization by solving jigsaw puzzles. Proc CVPR 2019, to appear. [2] R. PK. Poudel, S. Liwicki, R. Cipolla. Fast SCNN: Fast Semantic Segmentation Network. arXiv: 1902.04502v1, 2019 [3] E. Tartaglione, S. Lepsoy, A. fiandrotti, G. Francini. Learning sparse neural networks via sensitivity-driven regularization. Proc. NeurIPS 2018.

Skills and competencies for the development of the activity	The candidate is expected to have a Master of Science degree in Computer Engineering, Computer Science or similar fields, and prior knowledge in machine/deep learning, computer vision and statistical methods. Projects, tesi or publications in these areas will constitute a plus. The candidate is also expected to have programming experience in python