# AI and Computer Vision for Smart Cities

Giuseppe Amato*, Fabio Carrara*, Luca Ciampi*, Marco Di Benedetto*, Claudio Gennaro*,
Fabrizio Falchi*, Nicola Messina* and Claudio Vairo*
*Institute of Information Science and Technologies - National Research Council - Pisa, Italy

## I. Introduction

Artificial Intelligence (AI) is increasingly employed to develop public services that make life easier for citizens. In this abstract, we present some research topics and applications carried out by the Artificial Intelligence for Media and Humanities (AIMH) laboratory of the ISTI-CNR of Pisa about the study and development of AI-based services for Smart Cities dedicated to the interaction with the physical world through the analysis of images gathered from city cameras. Like no other sensing mechanism, networks of city cameras can 'observe' the world and simultaneously provide visual data to AI systems to extract relevant information and make/suggest decisions helping to solve many real-world problems. Specifically, we discuss some solutions in the context of smart mobility, parking monitoring, infrastructure management, and surveillance systems.

## II. Research Themes and Applications

### A. Visual Parking Lot Monitoring

Traffic-related issues are constantly increasing, and tomorrow's cities cannot be considered intelligent if they do not enable smart mobility. The AIMH laboratory proposes some Deep Learning (DL)-based solutions for parking lot monitoring running directly onboard embedded vision systems, i.e., devices equipped with limited computational capabilities that can capture and process images. Specifically, we introduced a decentralized solution for visual parking lot occupancy detection [1], which exploits Convolutional Neural Networks (CNNs) to classify the parking space occupancy directly onboard Raspberry Pi platforms equipped with camera modules. On the other hand, in [2] and [3], we extended this application by proposing a DL-based method that is instead able to estimate the number of vehicles present in the parking area without relying on meta-information regarding the monitored scene, such as the position of the parking lots. We show the output of our vehicle counting solution in Fig. 1. Moreover, unlike most of the works on this task, which focus on the analysis of *single* images, the AIMH group introduced, in [4], the use of multiple visual sources to monitor a wider parking area from different perspectives. The proposed multi-camera system is capable of automatically estimating the number of cars present in the *entire* parking lot directly on board the edge devices, combining an on-device DL-based detector and a decentralized geometric-based approach. Finally, in [5] we proposed a DL solution to detect and count vehicles in images taken from drones automatically.

This work was partially supported by H2020 project AI4media under GA 951911 and Extension (ESA, n. 4000132621/20/NL/AF).

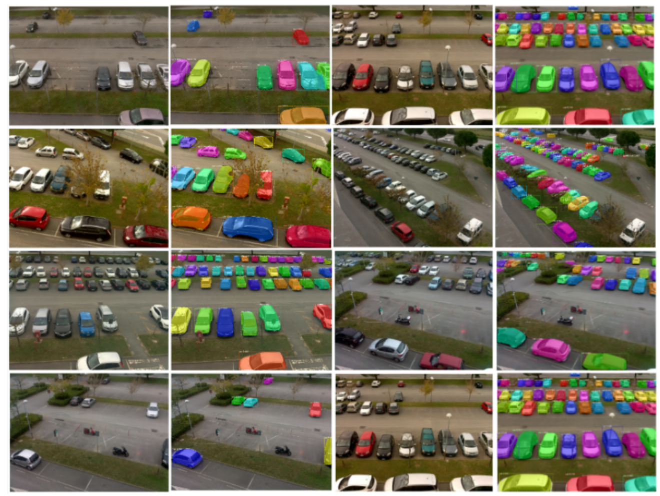

Fig. 1: **Examples of the output of our vehicle counting solution.** We show input images captured by smart cameras and the detected vehicles by our CNN-based technique.

### B. Visual People Detection and Counting

An essential task in many intelligent video surveillance systems is pedestrian detection since it is the main building block for many applications, such as people re-identification and counting. CNN-based pedestrian detectors have demonstrated their superiority over the approaches relying on hand-crafted features, but they require a considerable amount of labeled data for the learning stage. Since manually annotating new sets of images is expensive, an appealing solution is to gather synthetic data from virtual environments resembling the real world, where the labels are *automatically* collected by interacting with the graphical engine. In this direction, the AIMH group introduced *Virtual Pedestrian Dataset (ViPeD)* [6] [7], a new synthetic dataset generated with the highly photo-realistic graphical engine of a video game. We show a sample of this dataset in Fig. 2. We exploited it to train CNN-based pedestrian detectors. Moreover, in [8], we proposed a video-based counting technique that estimates the number of people present in the monitored scenes by taking advantage of the temporal correlation between consecutive frames.

### C. Visual Traffic Density Estimation

Monitoring vehicle flows in cities is crucial to improving citizens' urban environment and quality of life, and images

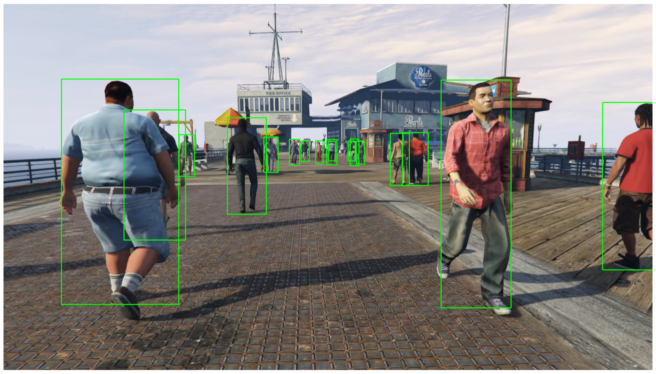

Fig. 2: **Sample of our *ViPeD* dataset**. Images and bounding boxes localizing the pedestrians are *automatically* gathered from a virtual-world. Image Courtesy of [7].

are the best sensing modality to perceive and assess the flow of vehicles in large areas. However, machine learning-based technologies hinge on large quantities of annotated data, preventing their scalability to city-scale as new cameras are added to the system. The AIMH laboratory proposes a technique that can automatically estimate the traffic density of urban scenarios by analyzing images [9] [10]. The main peculiarity of this methodology is that it can generalize to new sources of data for which there is no training data available. We achieved this generalization by exploiting an *Unsupervised Domain Adaptation (UDA)* strategy, whereby a discriminator attached to the output induces similar traffic density distribution in the test and train domains.

### D. Human Activity Monitoring

As occurs during a severe health emergency event such as the recent COVID-19 pandemic, there exist scenarios in which ensuring compliance to a set of guidelines becomes crucial to secure a safe living environment in which human activities can be conducted. The AIMH group proposes a deployed real use-case embedded system capable of perceiving people's behavior and aggregations, and able to supervise the appliance of a set of rules relying on a configurable plug-in framework [11]. As an effective setup, we implemented a set of visual-based modules for pedestrian detection, tracking, aggregation counting, social distancing calculations, and personal protection environment detection. To test the effectiveness of our solution, we monitored a known place in Italy during the restrictions imposed during the COVID-19 pandemic, proving satisfactory accuracy. In Fig. 3, we show an example of the functionality aiming at estimating the social distance.

## III. Conclusion

This abstract described several *human-centered AI* public services available to citizens proposed by the AIMH laboratory of the ISTI-CNR of Pisa. Specifically, we illustrated some AI applications based on the analysis of images gathered from city cameras that make life easier and safer for people in modern smart-cities, in the sphere of smart mobility, infrastructure management, and surveillance.

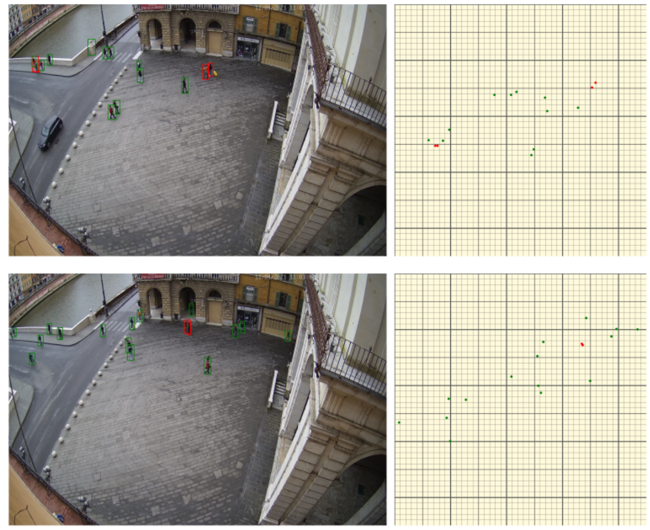

Fig. 3: **Examples of the output of our social distance measurer module.** *Left*: images with the detected pedestrians. *Right*: 2D projection on a virtual planar surface obtained through homography. Green color means a safe placement; red color indicates violations of the social distance rule. Image Courtesy of [11].

## References


[1] G. Amato, F. Carrara, F. Falchi, C. Gennaro, C. Meghini, and C. Vairo, "Deep learning for decentralized parking lot occupancy detection," *Expert Systems with Applications*, vol. 72, pp. 327–334, 2017.

[2] L. Ciampi, G. Amato, F. Falchi, C. Gennaro, and F. Rabitti, "Counting vehicles with cameras," in *Proceedings of the 26th Italian Symposium on Advanced Database Systems*, ser. CEUR Workshop Proceedings, vol. 2161. CEUR-WS.org, 2018.

[3] G. Amato, P. Bolettieri, D. Moroni, F. Carrara, L. Ciampi, G. Pieri, C. Gennaro, G. R. Leone, and C. Vairo, "A wireless smart camera network for parking monitoring," in *2018 IEEE Globecom Workshops*, dec 2018.

[4] L. Ciampi, C. Gennaro, F. Carrara, F. Falchi, C. Vairo, and G. Amato, "Multi-camera vehicle counting using edge-ai," *arXiv:2106.02842*, 2021.

[5] G. Amato, L. Ciampi, F. Falchi, and C. Gennaro, "Counting vehicles with deep learning in onboard UAV imagery," in *2019 IEEE Symposium on Computers and Communications (ISCC)*. IEEE, jun 2019.

[6] G. Amato, L. Ciampi, F. Falchi, C. Gennaro, and N. Messina, "Learning pedestrian detection from virtual worlds," in *International Conference on Image Analysis and Processing (ICIAP)*. Springer, 2019, pp. 302–312.

[7] L. Ciampi, N. Messina, F. Falchi, C. Gennaro, and G. Amato, "Virtual to real adaptation of pedestrian detectors," *Sensors*, vol. 20, no. 18, p. 5250, sep 2020.

[8] M. Avvenuti, M. Bongiovanni, L. Ciampi, F. Falchi, C. Gennaro, and N. Messina, "A spatio-temporal attentive network for video-based crowd counting," *CoRR*, vol. abs/2208.11339, 2022.

[9] L. Ciampi, C. Santiago, J. P. Costeira, C. Gennaro, and G. Amato, "Unsupervised vehicle counting via multiple camera domain adaptation," in *Proceedings of the First International Workshop on New Foundations for Human-Centered AI (NeHuAI)*, 2020, pp. 82–85.

[10] L. Ciampi, C. Santiago, J. Costeira, C. Gennaro, and G. Amato, "Domain adaptation for traffic density estimation," in *Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 2021.

[11] M. D. Benedetto, F. Carrara, L. Ciampi, F. Falchi, C. Gennaro, and G. Amato, "An embedded toolset for human activity monitoring in critical environments," *Expert Systems with Applications*, vol. 199, p. 117125, aug 2022.