Parking Lot Monitoring with Smart Cameras

Giuseppe Amato* Paolo Bolettieri* Fabio Carrara* Luca Ciampi* Claudio Gennaro* Giuseppe Riccardo Leone* Davide Moroni* Gabriele Pieri* Claudio Vairo*

***Abstract*—In this article, we present a scenario for monitoring the occupancy of parking spaces in the historical city of Lucca (Italy) based on the use of intelligent cameras and the most modern technologies of artificial intelligence.**

The system is designed to use different smart-camera prototypes: where the connection to the power grid is available, we propose a powerful embedded hardware solution that exploits a Deep Neural Network. Otherwise, a fully autonomous energy-harvesting node based on a low-energy custom board employing lightweight image analysis algorithms is considered.

I. INTRODUCTION

About 30% of rush hour traffic is generated by cars searching for a parking lot [1]. This is a common issue in every big or medium size city. These cars increase the congestion and the emission of noxious gases in the air; furthermore, the long amount of this wasted time is a cause of stress and late arrivals. An automatic system which monitors the urban streets could be very effective in addressing this problem by pointing on a map the areas in which parking spaces are easier to be found. Very often inductive field sensors installed under the street or one to one infrared sensors in indoor parking space are used to keep track of the status of the lots. These technologies become highly ineffective due to the high costs whenever large outdoor parking areas should be monitored. In addition, they require major and invasive work to be installed in existing parking areas and curb spaces, which are typical of historical cities. To address these issues, the research project SmartPark@Lucca aims at the development of a system based on a network of "intelligent" cameras able to monitor one or more parking areas of variable dimensions located in the medieval city of Lucca, Italy. Processing of the pictures captured by the cameras is performed on board of the smart cameras themselves so that no transmission of the pictures is required. Each camera is able to evaluate with extreme precision the occupancy level of the monitored parking area and to identify free or occupied spaces. Given that a single camera is able to monitor dozens of spaces, this infrastructure has a significantly lower cost, compared to other solutions. Inside the medieval city of Lucca there are places in which the wired electrical network is sometimes not present and it is not possible to put new artifact due to cultural heritage constraints. In order to manage these scenarios, our research project envisages the development of two solutions: a *Custom vision board* and a *Powered vision board*. The former solution is based on an embedded architecture that has low consumption and that allows us to use a battery pack and a module for harvesting energy through a photo-voltaic panel. The latter one is based on more commercial but more powerful solutions available off-the-shelf.

*All the authors are with the Institute of Information Science and Technologies (ISTI) of the National Research Council of Italy (CNR), Pisa, Italy, Name.Surname at isti.cnr.it

II. THE SOLUTIONS

The *ad hoc* realization of the smart camera prototype started from a deep study on the design of the architectural side. The guiding principle has been to be able to use state of the art computer vision technology using, at the same time, low cost sensors and electronic components, so that, once engineered, the device can be manufactured at low-cost in large quantities. The choice of processing unit and of the imaging sensor appears to be of particular relevance in this sense.

Another design requirement is represented by the ease of installation of the device, thus, the protective shield that has been considered for the sensor node is compact but able to accommodate all the components of the device.

The last and optional requirement is represented by the possibility to have a completely untethered device which communicates via wireless technologies and which has no need of electric power from the main supply. In this case the system has to have very low power consumption.

The design of the sensor node architecture is depicted in Fig. 1. Each single sensor node is composed of two main units. The first one is the Sensor Unit, devoted to the acquisition and processing of the images, and to the communication of the results to the system users. The second one is the (optional) Energy Harvesting Unit which serves as a regulator for battery charging and allows an optimal choice in terms of energy savings policies.

As mentioned above, in order to achieve all the goals we used two different types of processing boards: the powerful Nvidia Jetson[1] when the node is connected to the grid and

[1] https://developer.nvidia.com/embedded/buy/jetson-tx2

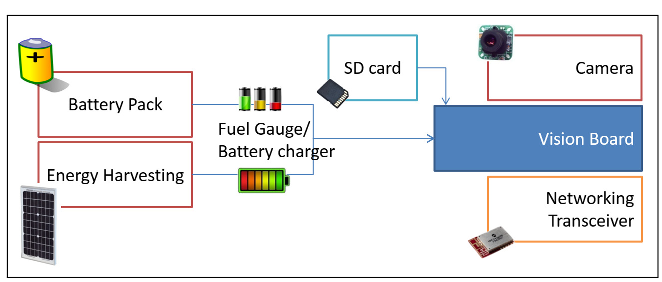

Fig. 1. Design of the architecture of the sensor node

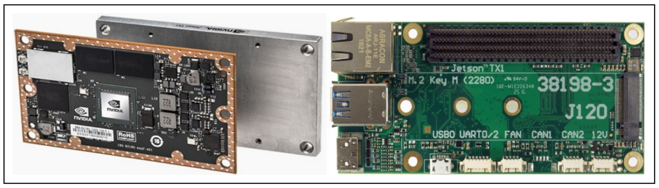

Fig. 2. Nvidia Tx1 processor (on the left) is mounted on the custom lightweight Auvidea J120 board (on the right)

a very low-powered custom board for the energy harvesting scenario.

A. Powered vision board based on Deep Learning

The embedded processor is the NVIDIA Jetson TX2, an ARM processor coupled with a GPU. With this equipment, we developed two approaches based on deep learning: the first solution is used to monitor the occupancy status of individual parking spaces in a parking lot. The second one is used to globally count cars parked in the parking lot. In the first scenario, we used *mAlexNet* [2], [3], a modified version of the *AlexNet* Convolutional Neural Network (CNN), in order to perform the parking space classification. This network is composed of three convolutional layers and two fully-connected layers and has been trained on the *CNRPark-EXT* dataset [2], a collection of roughly 150,000 bounding-box annotated images of vacant and occupied parking slots (called *patches*) in the campus of the National Research Council (CNR) describing most of the difficult situations that can be found in a real scenario: the images are captured by nine different cameras under various weather conditions, angles of view and light conditions.

In the second solution, we proposed a deep learning-based approach that is able to count the cars in captured images *without* any extra information of the scenes, like the regions of interest (i.e. the position of the parking lots) or the perspective map. This is a key feature since in this way our solution is directly applicable in unconstrained contexts. The proposed approach is based on *Mask R-CNN* [4], a CNN employed in many detection systems. Unlike previous methods that tackle the localization problem by building a sliding-window detector, *Mask R-CNN* solves the problem by operating within the recognition using regions paradigm, taking a full image as input and producing as output labels for each detected object together with bounding boxes and masks localizing them [5]–[7]. A similar approach has been exploited in a scenario of pedestrian detection [8].

This approach exploits a network of smart cameras (computing boards equipped with a camera module for acquiring images) where each smart camera monitors the set of parking spaces included in the portion of the parking lot seen by the camera. The system also comprises a server that receives the occupancy information of the parking spaces and visualizes them.

B. Custom vision board

For the realization of the custom vision board, it has been decided to realize a custom vision component by designing, printing and producing a new printed circuit board (PCB), since all existing prototyping boards have as common disadvantages high power consumption. The new PCB has been designed to have the maximum flexibility of use while maximizing the performance/consumption ratio. A good compromise has been achieved by using a *Freescale CPU* based on the ARM architecture, with support for MMU-like operating systems GNU/Linux. The chosen architecture has been proved to have an average consumption measured at the highest speed (454MHz) less than $500mW$.

The low powered custom board is not suitable for the deep learning approach. Specific artificial vision algorithms have been studied, designed and deployed for this board. The application defines some Regions of Interest (ROI) in the source image and uses a background model computed through lightweight methods [9]. An adaptive background is computed because it proved to be the most robust for use in uncontrolled outdoor scenes. The background is continuously updated using both the previous background model and the latest acquired actual image. In order to reach a strong belief of the status of each the parking slot, we perform two different image analyses, considering that an empty slot should appear as plain asphalt without nothing inside.

The first is the so-called “asphalt detection”: periodically checking small rectangular asphalt samples on the driveway (using the current background image so that no moving vehicle is on the region of interest) we identify similar hue and saturation values in the rest of the image.

For each ROI R_k the index $a_k(t)$ is computed. This index is proportional to the ratio of asphalt pixels with respect to the total number of pixels in the R_k. Then a very neat image of the contours of the vehicles is obtained with a Canny edge detection of the current foreground image. Similar to the first one the computed index $e_k(t)$ is proportional to the ratio of edge pixels in ROI R_k with respect to the total number of pixels in R_k. The combination of the two indexes creates the final belief of the sensor, which indicates the probability of the occupation of the parking lot.

$$P_k(t) = e_k(t) \cdot (1 - a_k(t)) \tag{1}$$

The occupancy status becomes effective only after being observed consecutively for a specific number of acquired frames.

III. Experiments

In order to test and compare our solutions, we have collected a new dataset acquiring images from a parking area located in Lucca. It consists of 3,890 images taken from two different points of views at different times of the day. We have randomly selected and labeled 100 of these images. The performance evaluators that we have used are Accuracy for the lot occupancy task and Mean Absolute Error (MAE) for the counting task. Results are shown in Table I.

TABLE I
RESULTS

Solution	Accuracy	MAE
mAlex	96%	-
Custom Vision Board	96%	-
Counting	-	3.6

IV. CONCLUSION

This research was carried out in the framework of the SmartPark@Lucca project, which aims at the development of a system based on a network of "intelligent" cameras able to monitor one or more parking areas of variable dimensions in the medieval city of Lucca, Italy. Given that a single camera is able to monitor dozens of spaces, this infrastructure has a significantly lower cost, compared to other solutions. We presented the work done in the first year of the project: the design and realization of a customizable smart camera prototype and the computer vision logic to evaluate the occupancy level of the monitored parking area and to identify free or occupied spaces. Due to cultural heritage constraints of the medieval historical center of Lucca sometimes the wired electrical network is not present: therefore we have also presented a lighter solution based on a custom vision board able to run using an energy harvesting unit. The prototype can host a powerful vision board to run state of the art CNN which achieve an error rate of 0.4%. Due to cultural heritage constraints of the medieval historical center of Lucca sometimes the wired electrical network is not present: in this case an energy harvesting unit in combination with a custom design very low consumption embedded vision board is able to run lightweight image processing with a slightly higher error rate of 0.65%. In both cases the processing of the pictures captured by the cameras is performed on board of the smart cameras themselves, so that no transmission of the pictures is required. We presented also the very first result of collaborative sensing that show how the overlapping of two cameras improve the monitoring in case of fixed or temporal occlusion. Cooperative sensing with three or more cameras will be the main focus of the next, second, year of the research project.

REFERENCES

[1] C. Dowling, T. Fiez, L. Ratliff, and B. Zhang, "Optimizing curbside parking resources subject to congestion constraints," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 5080–5085.

[2] 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.

[3] G. Amato, F. Carrara, F. Falchi, C. Gennaro, and C. Vairo, "Car parking occupancy detection using smart camera networks and deep learning," in *2016 IEEE Symposium on Computers and Communication (ISCC)*. IEEE, 2016, pp. 1212–1217.

[4] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.

[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 Communication (ISCC)*. IEEE, 2019.

[6] 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 (GC Wkshps)*. IEEE, 2018, pp. 1–6.

[7] L. Ciampi, G. Amato, F. Falchi, C. Gennaro, and F. Rabitti, "Counting vehicles with cameras." in *SEBD*, 2018.

[8] 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*. Springer, 2019, pp. 302–312.

[9] K. Kim, T. H. Chalidabhongse, D. Harwood, and L. Davis, "Background modeling and subtraction by codebook construction," in *2004 International Conference on Image Processing, 2004. ICIP'04.*, vol. 5. IEEE, 2004, pp. 3061–3064.