Deep Learning based video analytics solutions running on board of smart cameras

Vincenzo Carletti *†, Antonio Greco *†, Alessia Saggese *†, Mario Vento *†
* University of Salerno, Italy
† A.I. Tech - www.aitech.vision
Email: {carletti, greco, saggese, vento}@aitech.vision

***Abstract*—A.I. Tech is a spinoff company of the University of Salerno, designing and developing cutting edge video analytics solutions based on deep learning able to run directly on board of smart cameras, equipped with limited resources capabilities. A.I. Tech solutions that will be presented during the demo help to manage several vertical markets, from retail and business intelligence to security and safety, from smart parking to smart city and smart roads.**

I. Overview of the solution

Most of the deep learning based systems available nowadays in the market are based on the usage of off-the-shelf detectors (such as Yolo). Anyway, it is definitively more challenging to design and develop a system able to achieve comparable performance in terms of accuracy, but without the computational burden typically required by deep neural networks. This is especially true when the application has to run on board of smart cameras [1], having limited resources capabilities. Within this context, a common designing choice of all the A.I.Tech applications is to achieve an accuracy comparable with state of the art *heavy* neural network detectors and classifiers, but with a reduced hardware requirement and with a high frame rate capability. Thanks to this, A.I. Tech plugins are able to run directly on board of a huge amount of different smart cameras providing open platforms to specific partners (and in particular on board of specific models of the following camera manufacturers: Axis, Androvideo, Bosch, Dahua, Hanwha Techwin, Hikvision, Mobotix, Panasonic, Vivotek, Topview). A.I. Tech confirms to be in the world the video analytics vendor supporting the highest number of camera platforms. During the demo, we will show some of our applications, detailed in the following, running on board of a smart camera.

II. Video analytics products

In this section we are going to the describe the products that will be shown during the demo. More info are available in our website: www.aitech.vision

AI-RETAIL-DEEP is the video analytics solution for persons monitoring. Thanks to the combination between a proprietary deep learning based detector, a multi object tracker and a calibration mechanism, it is in charge of estimating the number of people inside an area; generating an alarm in case of overcrowding situations or in case of gathering detected; generating an alarm in case the social distances between two or more persons are not respected for a given amount of time; carrying out the counting of people by verifying the crossing of a line. An example of the solution in action is shown in Figure 1a.

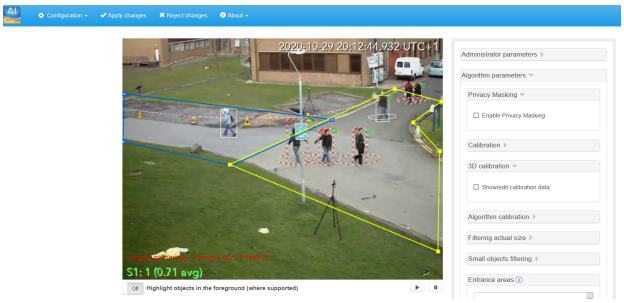
AI-PARKING-DEEP is able to automatically analyze a parking, so as to verify whether a place is free or occupied; it is a very useful product for monitoring a parking, both indoor and outdoor. The main advantage of this product lies in the fact that in order to monitor a spot, it is required that only a part of the vehicle is visible, without the need of seeing the whole vehicle while occupying a spot. An example of AI-PARKING in action is available in Figure 1c.

AI-TRAFFIC-DEEP is the video analytics solution for roads monitoring. It is in charge of identifying events of interest generated by vehicles and persons moving on the roads, to be used for both statistical and alarmistic purposes. Based on deep learning for the detection, on a multiobject tracking and on an advanced reconstruction in 3D of the scene, it allows the counting and classification of vehicles (car-motorcycle-truck), for each vehicle it performs the calculation of average speed and color evaluation;it evaluates the density of vehicles, raises an alarm in case a congestion is detected; furthermore, it allows to detect vehicles moving in the wrong direction or stopped in some forbidden areas and to detect the presence of pedestrians on the road. An example is reported in Figure 1b.

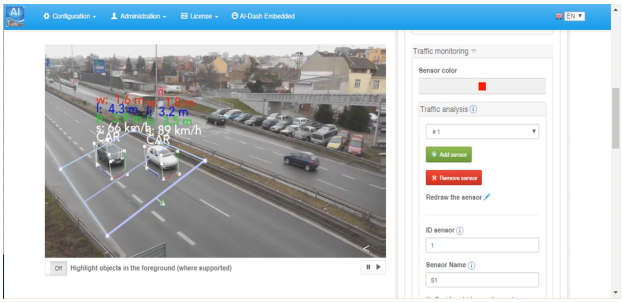
AI-LPR-DEEP: it is in charge of license plate detection and recognition. Differently from the other products available in the market, it is fully based on deep learning for both plate detection and license character recognition. An example of the product is shown in Figure 1d.

AI-BIO-DEEP: this is the plugin in charge of face analysis, aiming at detecting and evaluating the faces of the persons so as to extract both temporary and stable traits, such as gender [4], age, sentiment and ethnicity. AI-BIO-DEEP employs a deep learning based detector, especially devised for running on board of embedded platforms and smart cameras, and a multitasking deep learning based architecture for gender, age, ethnicity and sentiment analysis. This product can be used for both business intelligence applications and digital signage ones [3], aiming at personalizing advertisement contents shown on a monitor, depending on the specific person looking at the monitor. An example is shown in Figure 1e.

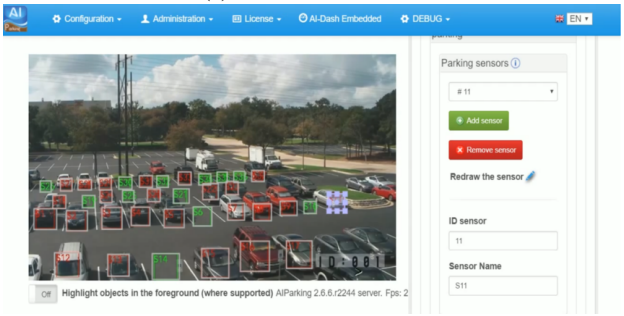
AI-SPILL-DEEP verifies if a person is falling down, by combining an advanced neural network for person classifica-



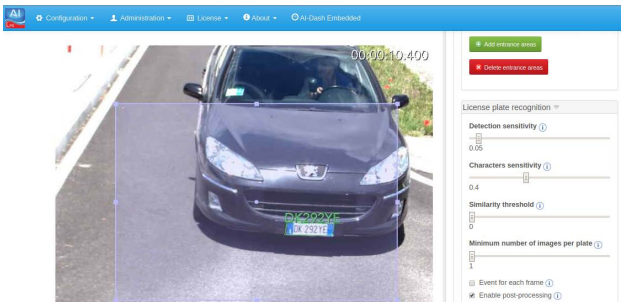
(a) AI-RETAIL-DEEP



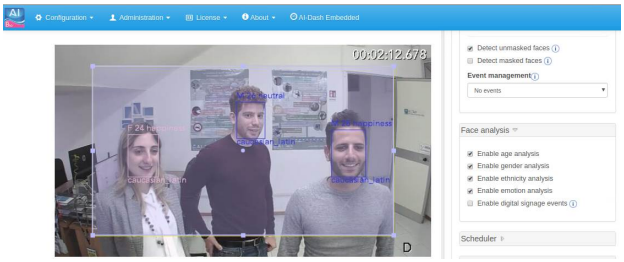
(b) AI-TRAFFIC-DEEP



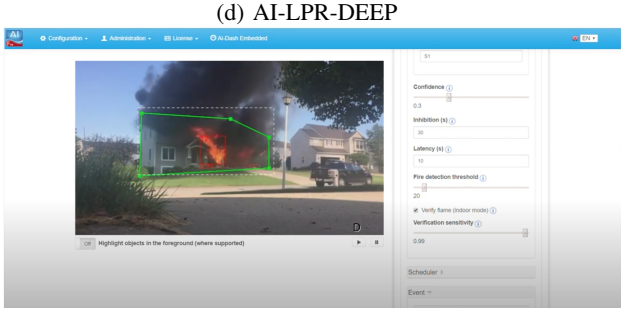
(c) AI-PARKING-DEEP



(d) AI-LPR-DEEP



(e) AI-BIO-DEEP



(f) AI-FIRE-DEEP

Fig. 1: Some example of A.I. Tech video analytics plugins in action. Fig. 1a AI-RETAIL-DEEP: the yellow area identifies the region where the analysis is performed. The dotted white-red bounding box around the persons identifies that that the social distances between that persons was not respected. Fig. 1b AI-TRAFFIC-DEEP: the area of interest where the evaluation is performed is in violet. A three dimensional bounding box is associated to each vehicle, together with the dimension of each object (w, l, h), expressed in meters; the speed (s), expressed in km/h; the category of the vehicle (*Car* in the example). Fig. 1c AI-PARKING-DEEP: the red boxes identifies occupied places, while green boxes identify available places. Fig. 1d AI-LPR-DEEP: in green we can see the license plate numbers detected by the application. Fig. 1e AI-BIO-DEEP: for each person, the rectangle of the face is reported and is colored in pink or in blue, depending on the gender of the person. For each face, the info about the sentiment, the age and the ethnicity are also reported. Fig. 1f AI-FIRE-DEEP: in green the area of interest, while the red box identifies the flame detected.

tion optimized for running on board of the camera with a mathematical model verifying the way in which the person moves, and in the specific case how the person falls down.

AI-FACEDETECT-DEEP: based on a single shot detector, the application is in charge of veryfying if a person is wearing or not a mask.

AI-FIRE-DEEP and AI-SMOKE-DEEP: they are in charge of detecting the flames and the smokes, without the need of using a thermal or a thermografic camera, but instead by means of a traditional optic sensor. Thanks to the employing of a deep neural network, it combines information based on the color, on the texture and on the motion [2].

References

[1] V. Carletti, A. Greco, A. Saggese, and M. Vento, "An effective real time gender recognition system for smart cameras," *J. Ambient Intell. Humaniz. Comput.*, vol. 11, no. 6, pp. 2407–2419, 2020.

[2] P. Foggia, A. Saggese, and M. Vento, "Real-time fire detection for video-surveillance applications using a combination of experts based on color, shape, and motion," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 9, pp. 1545–1556, 2015.

[3] A. Greco, A. Saggese, and M. Vento, "Digital signage by real-time gender recognition from face images," in *2020 IEEE International Workshop on Metrology for Industry 4.0 IoT*, 2020, pp. 309–313.

[4] A. Greco, A. Saggese, M. Vento, and V. Vigilante, "A convolutional neural network for gender recognition optimizing the accuracy/speed tradeoff," *IEEE Access*, vol. 8, pp. 130 771–130 781, 2020.