

# Cutting edge video analytics solutions: from the research to the market

Mattia Marseglia<sup>1,†</sup>, Domenico Rocco<sup>1,†</sup>, Stefano Saldutti<sup>1,†</sup> and Bruno Vento<sup>1,†</sup>

<sup>1</sup>A.I. Tech srl - [www.aitech.vision](http://www.aitech.vision), Piazza Vittorio Emanuele 10, Penta(SA), 84123, Italy

## Abstract

A.I. Tech was born as a spinoff company of the University of Salerno and designs and develops cutting edge video analytics solutions based on deep learning, able to run on board of smart cameras and/or on devices with limited resource capabilities. A.I. Tech solutions are designed to serve various vertical markets: retail, business intelligence, security and safety, smart parking, smart city and smart roads. In this paper we present all these solutions, which are the products of years of research transferred to the market.

## Keywords

A.I. Tech, video analytics, cutting edge, computer vision

## 1. Company presentation

A.I. Tech designs and develops cutting edge video analytics solutions based on the most advanced artificial intelligence and deep learning algorithms, also running directly on board of smart cameras, and therefore optimized for low-performance hardware. A.I. Tech boasts partnerships with world leaders in their reference fields, including (the list is not exhaustive) NVIDIA, Panasonic, Samsung, Hanwha Techwin, Mobotix, Axis, Hikvision, Dahua. In particular, Hanwha Techwin, Panasonic and Mobotix resell the video analytics solutions from A.I. Tech on a global scale. In 2017 A.I. Tech has been selected among the Top25 international companies in the field of Artificial Intelligence by CIO Applications Magazine. In 2018 it enters the Top10 Most Innovative AI Solution Providers. Its technology was selected among the finalists in the Benchmark Innovation Award in 2018, 2019, 2020, 2021 and 2022. In 2018 it wins the award in the Business Intelligence category, with the AI-RETAIL video analytics solution. In 2020 A.I. Tech won the Corporate LiveWire award in the “Most Innovative in Video Analytics” category. In 2020 its solutions are finalists in the Security and Fire Excellence Award, for the AI-CROWD-DEEP product (in the Security Software Product Innovation of the Year category) and for the WOW project (in the Security Project of the Year category). The AI-TRAFFIC solution for traffic monitoring is also the winner of the IoMOBILITY AWARD 2020, in the Mobility Analytics category. Corporate LiveWire awarded A.I.

Tech the “Innovation & Excellence Awards” for the year 2022, renewing the award also for the year 2023, considering the company as the most innovative in the field of “AI Technology”.

The activities that A.I. Tech carries out, with a highly technological and scientific content, require specialized skills in the field of Artificial Intelligence, Artificial Vision and Embedded Systems. For this reason, the company has a very close collaboration relationship with the Department of Information and Electrical Engineering and Applied Mathematics (DIEM) of the University of Salerno. In particular, there is also an agreement for the activation of company internships as well as scientific collaborations for the next years. These activities allow to transfer the scientific skills of the DIEM research group in the field of Artificial Vision and Artificial Intelligence, with a consequent technological transfer of research products which takes the form of a series of cutting edge artificial intelligence products, commercially available at an international level.


## 2. Overview of the solutions

Most of the deep learning based systems available nowadays in the market are realized on top of off-the-shelf detectors. Anyway, designing software solutions engineered to be as accurate as the state-of-the-art without the computational burden typically required by deep neural networks, is definitively more challenging. Realizing computationally inexpensive solutions is a mandatory requirement in several real-world applications where the system is expected to process hundreds of video streams simultaneously in real-time keeping an affordable cost; smart-cities are a noteworthy example of that. Moreover, in different contexts the processing is required to be performed on the edge due to environmental constraints, therefore the video analytic application has to run on

*Ital-IA 2024: 4th National Conference on Artificial Intelligence, organized by CINI, May 29-30, 2024, Naples, Italy*

<sup>†</sup>These authors contributed equally.

✉ [mattia.marseglia@aitech.vision](mailto:mattia.marseglia@aitech.vision) (M. Marseglia);  
[domenico.rocco@aitech.vision](mailto:domenico.rocco@aitech.vision) (D. Rocco);  
[stefano.saldutti@aitech.vision](mailto:stefano.saldutti@aitech.vision) (S. Saldutti); [br1.vento@gmail.com](mailto:br1.vento@gmail.com)  
(B. Vento)

 © 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

board of smart cameras [1], with very limited hardware resources.

Within this context, a common design choice of all the A.I.Tech applications is to preserve the accuracy comparable with state-of-the-art detectors and classifiers based on *heavy* neural networks, but achieving the lowest hardware requirement together with the higher processing throughput. 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: Androvideo, Axis, Bosch, Dahua, Hanwha Techwin, Hikvision, Mobotix, Panasonic, Topview, Vivotek). A.I. Tech confirms to be, in the world, the video analytics vendor supporting the highest number of camera platforms.

### 3. Video analytics products

In this section we are going to describe 12 video analytics solutions currently available on the market.

AI-BIO<sup>1</sup> performs face analysis with the purpose of extracting soft-biometric features like age, gender and emotion [2, 3, 4]. The application has a multitask architecture based on multiple deep neural networks engineered to be executed on board of embedded platforms and smart cameras. It can be used both for business intelligence and for digital signage applications [5]. In particular, in the last case, the aim is to personalize advertisement contents on a monitor by taking into account the soft-biometric features extracted from the face of the person who is watching at the monitor. An example is shown in Figure 1a.

AI-CROWDCOUNTING<sup>2</sup> is a video analytics application tailored to estimate, for statistical or alerting purposes, the crowd density within specific very crowded areas of interest. Powered by a deep learning model and boosted by a distinctive training strategy [6], the system is not only able to detect people fully visible in the scene, but also to identify those that are very occluded, thanks to a point-based head detection algorithm. This makes the application particularly suited for very crowded environments, such as stadiums, concerts or trade fairs. Figure 1b shows an example of the solution in action.

AI-CROWD-DEEP<sup>3</sup> is the video analytic solution for people monitoring. Thanks to the combination of a proprietary deep learning based detector, a multi object tracker [7] and a calibration mechanism, it is capable of: (i) estimating the number of people inside an area; (ii) generating an alarm in case of overcrowding situations or in case of gathering detected; (iii) generating

an alarm if two or more persons are not respecting the social distances for a given amount of time; (iv) counting of people that cross virtual lines; (v) counting the number of pedestrians crossing one area and arriving in another, building the origin-destination matrix. An example of the solution in action is shown in Figure 1c.

AI-FIREPLUS<sup>4</sup> are the solutions focused on the early detection of fires. It combines the analysis of movement and appearance with a deep neural network to detect the presence of flame or smoke within an area under monitoring [8], it can operate in both indoor and outdoor environments. The main benefit of this application is that it does not require thermal or thermographic sensors, but traditional optic ones instead. An example is shown in Figure 1d.

AI-INTRUSION<sup>5</sup> is the video analytic solution for the detection of intruders (people or vehicles). It is capable to detect: (i) intrusions or loitering within an area of interest framed by the camera; (ii) the crossing of a virtual line; (iii) the crossing of multiple crossing lines (not necessarily parallel) in sequence. In addition to the size and the aspect ratio of the object, it uses a deep neural network to filter objects according to their class. An example is reported in Figure 1e.

AI-LOST<sup>6</sup> is the video analysis application designed to detect removed or abandoned objects in restricted environments where constant surveillance cannot be guaranteed [9]. The application can use a deep neural network to recognize garbage or, alternatively, baggage. An example is reported in Figure 1f.

AI-LPR is the solution for license plate detection and recognition. Unlike 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 1g.

AI-PARKING<sup>7</sup> is designed to monitor both indoor and outdoor parking, so as to verify whether a parking spot is free or occupied. Unlike other solutions based on vehicle detection, this is a very effective application requiring that only a part of the vehicle must be visible to monitor a spot. An example of AI-PARKING in action is available in Figure 1h.

AI-PEOPLE-DEEP<sup>8</sup> is the solution that exploits a deep neural network to count the people framed by a camera positioned in zenithal view. Inspired by [10], the application is designed to work both indoors and outdoors where it is possible to ensure that the illumination conditions controlled. An example is reported in Figure 1i.

AI-PPE<sup>9</sup> is designed to detect people wearing personal

<sup>1</sup><https://www.youtube.com/watch?v=awze1fHoQEE>

<sup>2</sup><https://youtu.be/h0qDXkZkObU?si=Su6gStufv9NbUrK9>

<sup>3</sup><https://www.youtube.com/watch?v=BiCyon1KZco>

<sup>4</sup><https://www.youtube.com/watch?v=U1SwnESua0g>

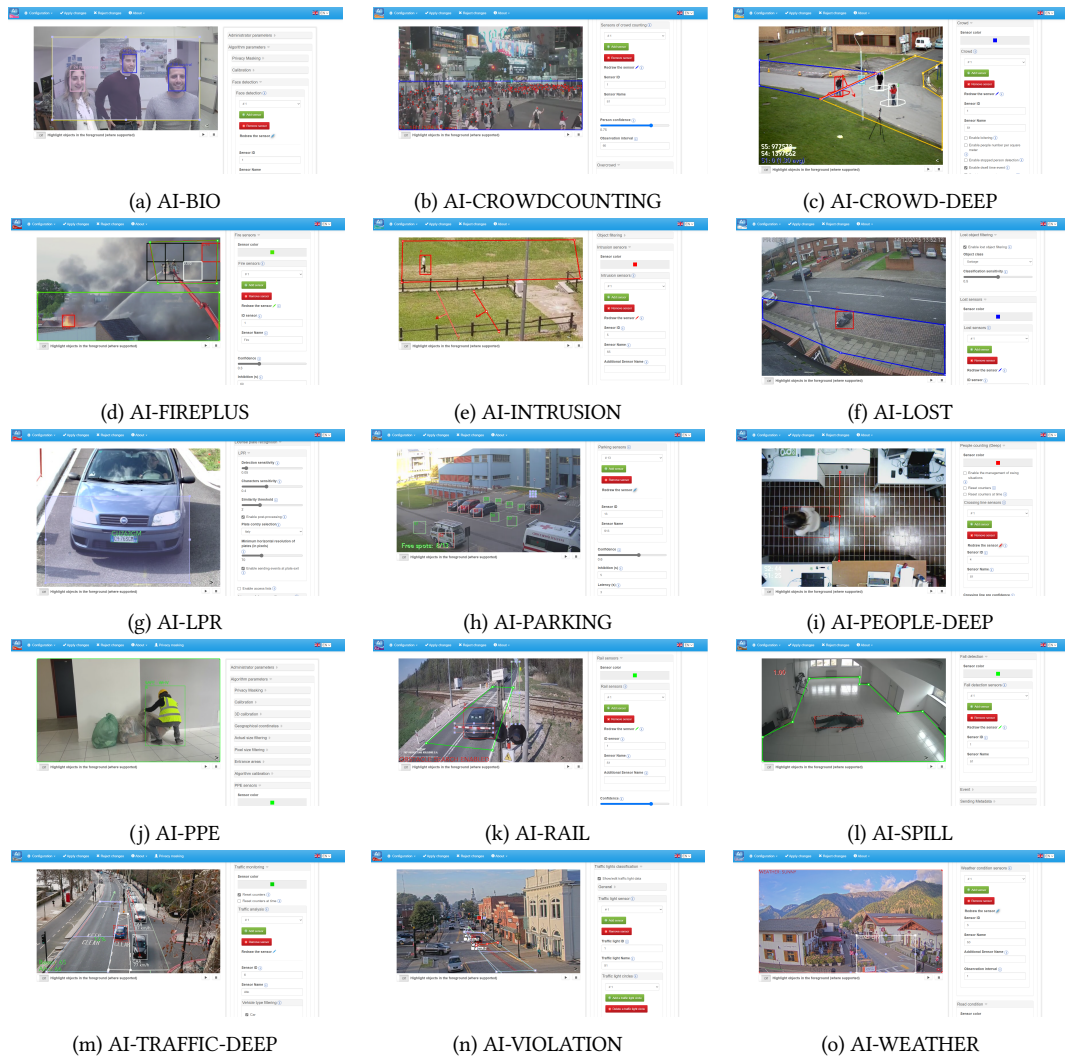
<sup>5</sup><https://www.youtube.com/watch?v=3kUUOcofVow>

<sup>6</sup><https://www.youtube.com/watch?v=gq24PrW6UwQ>

<sup>7</sup><https://www.youtube.com/watch?v=VDQ82Di4fZs>

<sup>8</sup>[https://www.youtube.com/watch?v=x6N5g4Fs6\\_U](https://www.youtube.com/watch?v=x6N5g4Fs6_U)

<sup>9</sup><https://www.youtube.com/watch?v=-fz25HYcFL0>



**Figure 1:** Some examples of A.I. Tech video analytic plugins in action. Fig. 1a AI-BIO: for each person, the rectangle around the face is shown in pink or in blue, depending on the gender of the person; moreover, the figure shows all the soft-biometric features extracted by the software: the emotion and the age. Fig. 1b AI-CROWDCOUNTING: for each detected person, the application draws a red point, showing in real-time the number of people present in the region of interest. Fig. 1c AI-CROWD-DEEP: the yellow area highlights the region where the analysis is performed. The dotted white-red bounding box around emphasizes a cluster of people that are not respecting the social distances. Fig. 1d AI-FIREPLUS: in green the area of interest, while the red box calls attention to the detected flame. In the black grid, the detected smoke is highlighted in red. Fig. 1e AI-INTRUSION: the intrusion area is the red polygon and the multiple crossing lines are the numbered red lines below. In the example, a person has been detected in the intrusion area. The P at the top left of the bounding box indicates that the object is a person (rather than V for vehicle). Fig. 1f AI-LOST: the area of interest is the polygon in blue. The red bounding box with the G string indicates that the detected object is garbage (instead of B for baggage). Fig. 1g AI-LPR: in green we can see the license plate numbers recognized by the application. Fig. 1h AI-PARKING: the red boxes highlight occupied spots, while green boxes those that are free. Fig. 1i AI-PEOPLE-DEEP: a red bounding box is drawn when a person crosses the virtual line. Fig. 1j AI-PPE: for each detected person, the application draws a bounding box and a string indicating the recognized tool (W for no ppe, WH for only helmet, WV for only vest and WHV for both helmet and vest). Fig. 1k AI-RAIL: a red bounding box is drawn around detected objects if they are within a restricted area when the barrier blocks the road. Fig. 1l AI-SPILL: in the scene a red bounding box is drawn around a person fallen within the area of interest. Fig. 1m 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 three dimensions of each object (width, length, height), expressed in meters; the speed (s), expressed in km/h; the category of the vehicle (Car in the example). Fig. 1n AI-VIOLATION: the status of the traffic light is shown in the box on the side (in the example), the area where the analysis is performed is in violet, the application allows to draw the limit of the stopping line (red line). Fig. 1o AI-WEATHER: A sensor is placed near the road for monitoring, and another sensor covering the entire image is utilized for classifying weather conditions. After the observation time within the sensors has passed, the classification outputs are displayed.

protective equipment (PPE). The application is based on the architecture described in [11]. The PPE combinations that the application is able to detect are: "Helmet", "Vest" and "Helmet and Vest". This solution can be used both in the case of access control system and for the surveillance of construction sites or places where works are in progress. In the first case, the use of the product is meant to verify that a worker is wearing the specified PPE, in order to authorize him to enter a work area. In the second, the product can be used for continuous monitoring of a work area with the aim of verifying that workers are wearing all the PPE required. An example of the product is reported in Figure 1j.

AI-RAIL<sup>10</sup> is a video analysis application designed for enhancing railway safety. It combines traditional computer vision techniques along with deep neural networks to identify and analyze the behavior of vehicles, pedestrians, and obstacles within sensitive areas such as level crossings area or along railway lines. The analysis can be activated depending on the barrier status, which can be obtained by either an external signal or through neural networks integrated into the system. An example is shown in Figure 1k.

AI-SPILL<sup>11</sup> is designed to monitor a person walking in an unsupervised area and detect if the person falls, raising an alarm if that happens. The analysis is performed using a mathematical model that allows to analyse the behavior of a person moving in the scenario of interest, especially walking and falling dynamics. An advanced neural network, trained with thousands of fallen people samples and optimized for running on board the camera, is then used to confirm the initial outcome of that model. An example is reported in Figure 1l.

AI-TRAFFIC-DEEP<sup>12</sup> is the video analysis solution for road monitoring for both statistical and alarmist purposes. Technically speaking, the application is based on a deep learning based vehicle and people detector [12] followed by a multi-object tracking module [7] and an advanced 3D scene reconstruction stage. It is capable of: (i) counting and classifying vehicles among cars, motorcycles and trucks; (ii) estimating the average speed and the color of each detected vehicle; (iii) evaluating the density of vehicles on a road branch and raise an alarm if congestion is detected; (iv) detecting vehicles travelling in the wrong direction or that stopped in some forbidden areas; (v) detecting the presence of pedestrians on the road; (vi) counting the number of vehicles and pedestrians crossing one area and arriving in another, building the origin-destination matrix; (vii) detecting lane changes and abnormal maneuvers (such as U-turns in prohibited areas) made by vehicles, based on crossing a set of user-configured virtual lines. An example is reported in Figure

<sup>10</sup>[https://youtu.be/cDh1epks3x0?si=TCZlm8QJOG\\_FJ6bk](https://youtu.be/cDh1epks3x0?si=TCZlm8QJOG_FJ6bk)

<sup>11</sup><https://www.youtube.com/watch?v=pCFBnWC8uPQ>

<sup>12</sup>[https://www.youtube.com/watch?v=6yQS6n\\_nTcl](https://www.youtube.com/watch?v=6yQS6n_nTcl)

1m.

AI-VIOLATION<sup>13</sup> is a vertical solution able to detect traffic light violations (see Fig. 1n), namely the presence of vehicles crossing the stopping line while the traffic light is red. It is based on the above mentioned vehicle detector and a classifier that allows surveillance cameras (which are commonly installed over the city) to read the traffic light status without the need to install external devices. The state of a traffic light includes the color of the active traffic light circle and whether it is blinking or not. In particular, the application can identify vehicles crossing the stop line at the traffic light while the traffic light status is red and send a notification to report the violation. This notification contains also information about the vehicle, such as the type (between motorcycle, bicycle, car, truck), the estimated average speed and all the information that are necessary to decide whether there are legal limits for a fine.

AI-WEATHER<sup>14</sup> is an innovative application that uses deep neural networks to monitor weather and road conditions. This app can recognize a wide range of weather states, including sunny, cloudy, rainy, snowy and foggy, as well as road surface conditions, which can vary between dry, non-dry and flooding. This application is designed to operate effectively in outdoor environments and requires visibility of both the road surface and the sky at the same time (see Fig. 1o). AI-Weather offers a variety of useful alerts to users, including sending periodic updates on weather and road conditions, as well as instant notifications when the status of one of the sensors changes.

## References

- [1] V. Carletti, A. Greco, A. Saggese, M. Vento, An effective real time gender recognition system for smart cameras, *J. Ambient Intell. Humaniz. Comput.* 11 (2020) 2407–2419.
- [2] A. Greco, A. Saggese, M. Vento, V. Vigilante, A convolutional neural network for gender recognition optimizing the accuracy/speed tradeoff, *IEEE Access* 8 (2020) 130771–130781. doi:10.1109/ACCESS.2020.3008793.
- [3] A. Greco, A. Saggese, M. Vento, V. Vigilante, Gender recognition in the wild: a robustness evaluation over corrupted images 12 (2021).
- [4] A. Greco, A. Saggese, M. Vento, V. Vigilante, Effective training of convolutional neural networks for age estimation based on knowledge distillation, *Neural Comput. Appl.* (2021).
- [5] A. Greco, A. Saggese, M. Vento, Digital signage by real-time gender recognition from face images, in:

<sup>13</sup><https://www.youtube.com/watch?v=gAVEHPCckbE>

<sup>14</sup>[https://www.youtube.com/watch?v=\\_gn-odtuWJo](https://www.youtube.com/watch?v=_gn-odtuWJo)

- 2020 IEEE International Workshop on Metrology for Industry 4.0 IoT, 2020, pp. 309–313.
- [6] L. Fotia, G. Percannella, A. Saggese, M. Vento, Highly crowd detection and counting based on curriculum learning, in: International Conference on Computer Analysis of Images and Patterns, Springer, 2023, pp. 13–22.
  - [7] P. Foggia, G. Percannella, A. Saggese, M. Vento, Real-time tracking of single people and groups simultaneously by contextual graph-based reasoning dealing complex occlusions, in: 2013 IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS), IEEE, 2013.
  - [8] P. Foggia, A. Saggese, 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* 25 (2015) 1545–1556. doi:10.1109/TCSVT.2015.2392531.
  - [9] P. Foggia, A. Greco, A. Saggese, M. Vento, A method for detecting long term left baggage based on heat map., in: *VISAPP* (2), 2015, pp. 385–391.
  - [10] A. Greco, A. Saggese, B. Vento, A robust and efficient overhead people counting system for retail applications, in: *International Conference on Image Analysis and Processing*, Springer, 2022, pp. 139–150.
  - [11] A. Greco, S. Saldutti, B. Vento, Fast and effective detection of personal protective equipment on smart cameras, in: *International Conference on Pattern Recognition*, Springer, 2022, pp. 95–108.
  - [12] A. Greco, A. Saggese, M. Vento, V. Vigilante, Vehicles detection for smart roads applications on board of smart cameras: A comparative analysis, *IEEE Trans. Intell. Transp. Syst.* (2021) 1–13.